COMPARATIVE PHYLOGEOGRAPHIC PATTERNS AMONG SELECTED INDIGENOUS AND INTRODUCED COLLEMBOLA ON MARION ISLAND

Marike Myburgh

Thesis presented in partial fulfilment of the requirements for the degree of Master of Science at the University of Stellenbosch.

Supervisors: Drs. Bettine Jansen van Vuuren and Savel R. Daniels

August 2005

Verklaring

Ek, die ondergetekende, verklaar hiermee dat die werk in hierdie tesis vervat, my eie oorspronklike werk is en dat ek dit nie vantevore in die geheel of gedeeltelik by enige universiteit ter verkryging van ‘n graad voorgelê het nie.

Declaration

I, the undersigned, hereby declare that the work contained in this thesis is my own original work and that I have not previously in its entirety or in part submitted it at any university for a degree.

Handtekening / *Signature*: .....

Datum / *Date*: .....

Acknowledgements

There are a number of institutions and people without whom this project would not have been possible:

The USAID (United States Agency for International Development) provided bursary support, and together with the CBP-CCR (Capacity Building Programme for Climate Change Research) provided funding. The Department of Environmental Affairs and Tourism: Antarctica and Islands are thanked for logistical support on both the 2003 and 2004 voyages to Marion Island.

I would like to give thanks to Professor Steven Chown for initiating this project and securing funding, and to numerous individuals of the Centre for Invasive Biology for useful comments and support in the completion of this project. Also, to my supervisors Bettine Jansen van Vuuren and Savel Daniels I would like to extend my gratitude.

My sincerest appreciation toward various members of the Marion 59 and Marion 60 teams as well as numerous others on the relief voyage to Marion in 2003, such as Richard Mercer, Sarette Slabber, Erika Nortjé, Jacques Deere, Peter le Roux, Lizel Hugo and Michelle Greve. I am also grateful to Anni Schwizer and Mark Keith for the collection of samples. A special thanks to the coordinator Mr Adriaan Dreyer who ensured that I managed to get around to all my sampling localities.

Thank you to the members of the Evolutionary Genomics Group at the University of Stellenbosch for many of cups of tea and helpful comments, especially from Sandy Willows-Munro, Liezl Mortimer and Dr Krystal Tolley. My heartfelt thanks also to Kenneth Oberholzer for help in the final stages of my thesis and to Tinus Kotzé, who spent a lot of time checking my formatting.

Most importantly, I would like to thank my friends and family who were always there when things were tough. It was a rough three years and I have grown tremendously as a person during this time, but it would not have been possible without all of you.

Thank you, one and all…

Opsomming

Marion Eiland is die groter van twee eilande wat die Prins Edward eilandgroep vorm, ongeveer 2600 km suid-oos vanaf Kaapstad. Gedurende glasiasie episodes in die verlede, was Marion gedeeltelik met ys bedek en fauna en flora was geisoleer in klein gedeeltes van die eiland wat beskut was. As gevolg van hierdie langtermyn isolasie word daar voorgestel dat spesies wat op die eiland was voor hierdie gebeurtenisse plaasgevind het, ‘n geografiese verdeling van genetiese variasie sal toon. Sedertdien het die meerderheid van die eilande in die Suidelike Oseaan ook ‘n geskiedenis van skepe wat op een of meer van die eilande aangedoen het in die jag op robbe. Hierdie besoeke deur die robjagters aan die eilande het meegebring dat daar eksotiese spesies op die eiland aangekom het of dat spesies wat reeds op die eiland voorgekom het, opnuut daar geland het. Hierdie, tesame met die onlangse wetenskaplike reise na Marion Eiland, het daartoe gelei dat baie eksotiese spesies ingevoer is na die eiland. Twee inheemse spesies, naamlik *Cryptopygus antarcticus* (Willem, 1901) en *Tullbergia bisetosa* (Börner, 1903) is gekies om as model spesies te dien om die moontlike gevolge van glasiasie op die geografiese verspreiding van genetiese variasie te toets. In ag geneem die moontlike negatiewe impak wat eksotiese spesies op die inheemse spesies mag hê, is die filogeografie van ‘n uitheemse spesie, *Isotomurus* cf. *palustris* (Müller, 1876) ook bepaal en vergelyk met díe van die bogenoemde inheemse spesies. Dit is gedoen deur twee gene, sitokroom oksidase I (COI) en sitokroom oksidase II (COII) te analiseer. Hierdie gene se basispaarvolgordes is bepaal (GENBANK toetreenommers DQ147289-DQ147558), en hulle is ge-analiseer deur gebruik te maak van AMOVA (Analise van Molekulêre Variasie), SAMOVA (Ruimtelike Analises van Molekulêre Variasie) en NCA (geseteldegroepsanalises). Net soos verwag, het die twee inheemse spesies albei duidelike tekens getoon van ‘n bevolkingstoename terwyl die uitheemse spesie amper geen variasie oor sy verspreidingsgebied getoon het nie. Die lewensstyle van spesies het ‘n beduidende invloed op hulle genetiese bevolkingstruktuur. As sulks is díe in ag geneem om die verskille tussen hulle genetiese strukture te verduidelik.

Abstract

Marion Island is situated approximately 2600 km southeast of Cape Town and is the larger of the two islands that comprise the Prince Edward Island group. During past glaciation events, Marion was partially covered by ice with fauna and flora confined to isolated refugia across the island. As a result of these long-term isolation events, it is postulated that species predating these glaciation events might show geographic partitioning of genetic variation. Subsequently, the majority of Southern Oceanic islands have a history of sealing activities with vessels frequenting various islands in their hunt for seals. These combined visits to several islands could have facilitated the introduction of alien species, or the reintroduction of species already present on islands. These incidents, combined with more recent scientific voyages, have led to the establishment of several exotic species on Marion Island. Two indigenous Collembola species (*Cryptopygus antarcticus* (Willem, 1901) and *Tullbergia bisetosa* (Börner, 1903)) were chosen as model species to investigate the possible effects of glaciation on the spatial distribution of genetic variation on indigenous species. Given the negative impact that alien taxa have on indigenous species and ecosystems, the phylogeographic population structure of the recently introduced *Isotomurus* cf. *palustris* (Müller, 1876) was determined and compared to those described for the two indigenous species. To address these questions, two mitochondrial DNA genes were targeted: cytochrome oxidase subunit I (COI) and cytochrome oxidase subunit II (COII). These genes were sequenced (GENBANK accession numbers DQ147289-DQ147558) and analysed using, amongst others, AMOVA (Analysis of Molecular Variance), SAMOVA (Spatial Analysis of Molecular Variance) and NCA (Nested Clade Analysis). As expected, the two indigenous species showed distinct signs of population expansion, whilst the recently introduced species exhibits little genetic variance across its range. The life histories of species have an influence on their resultant genetic signature; therefore life history parameters were incorporated to explain differences in the phylogeographic patterns observed for these three species.

Table of contents

List of figures

List of tables

Aims

The genetic structure of populations and species reflects the effective movement of individuals and their genes. In this study we attempt to explain how climate change and life history parameters have affected contemporary population genetic structure. We describe the genetic population structure of three selected Collembola species on Marion Island. In addition, we compare observed geographic partitioning of genetic variation among two indigenous and one recently introduced species of springtail on the sub-Antarctic Marion Island.

Specific objectives:

1. To describe the geographic partitioning of genetic variation for two indigenous (*Cryptopygus antarcticus* and *Tullbergia bisetosa*) and one recently introduced species (*Isotomurus* cf *palustris*) using two mitochondrial DNA protein coding genes, cytochrome oxidase subunit I (COI) and cytochrome oxidase subunit II (COII).

2. To compare the genetic signatures of indigenous and (recently) introduced species across Marion Island.

3. To explain whether / how different life histories influenced the phylogeographic structure of selected Collembola across Marion Island.

Chapter 1

Introduction

Introduction

Islands

When conspecific populations are isolated (either spatially or temporally) for extensive durations with limited or no gene flow, evolutionary processes such as random genetic drift and mutations will promote genetic divergence between them (Ridley, 1993; Gillespie & Roderick, 2002). Therefore, the more isolated a population is, the less gene flow there will be between it and the populations at the source of colonisation. This in turn causes populations to diverge from the source population due to random genetic drift as well as the accumulation of mutational changes over time (Ridley, 1993; Gillespie & Roderick, 2002; Edmonds *et al*, 2004). In a spatial context, gene flow may be obstructed by various barriers such as mountains, rivers or simply patches of unsuitable habitat (Ridley, 1993; Miller & Harley, 1996; Tamarin, 1999). A classical example would be taxa that inhabit oceanic islands where the surrounding water serves as a barrier to limit or prevent gene flow. Such island systems are particularly attractive environments to study the processes that sculpt evolution for various reasons: firstly they exist within definite boundaries, secondly, there is reduced gene flow between them, thirdly categorisation of fauna and flora is easier than on continental systems due to the smaller size of the islands, and lastly, they often show a diversity of habitats (Emerson, 2002).

Two main types of oceanic islands have been defined by Gillespie and Roderick (2002). First, "Darwinian" or new islands that have formed *de novo* through processes such as volcanism and secondly, "fragment" islands that were formed by continental fragmentation. In the latter case the same organisms that inhabit the continental source will already be present on the islands. Conversely, on newly formed islands all ecological niches will initially be vacant. Gillespie and Roderick (2002) argue that the composition of an island's fauna depends on the scale of its isolation. Numerous factors such as the nature of the surrounding matrix, other islands that can serve as stepping-stones from the source of colonisation and the availability of transport vectors determine the isolation of an island (Gillespie & Roderick, 2002).

The entire area from the coast of Antarctica to the Antarctic Polar Frontal Zone (APFZ) contains only a few isolated island groups (Hall, 2002), the majority being of volcanic origin (Hall, 1990). The Southern Ocean has been divided into three different geological provinces. The South Indian Ocean Province (formerly the Kerguelen Province) includes the Prince Edward Islands, Crozet, Kerguelen, Heard and MacDonald Islands (Hänel & Chown, 1998) and stretches from the coast of mainland Antarctica (Enderby) to 45ºS and between 30ºE and 90ºE into the Southern Indian Ocean (Pugh, 1993).

The Prince Edward Island group is comprised of two islands, separated by 19 km of ocean: Marion Island (46º54'S, 37º45'E) and Prince Edward Island (46˚38'S, 37˚57'E). Both islands are peaks of a submerged volcano (Hänel & Chown, 1998; Hall, 1990; 2002). These islands are considered semi-closed systems where the terrestrial ecosystems are vastly different from the surrounding ocean and yet are dependent on it. When "Darwinian" islands, such as these former two volcanic islands, remain isolated from their source of colonisation for prolonged periods, evolutionary processes will lead to diversification and the formation of neo-endemic species. This in turn will result in unique biological communities being established on these islands (Gillespie & Roderick, 2002). It is therefore not surprising that sub-Antarctic islands have many terrestrial hexapod and plant species that are endemic to one or a few groups of islands (Bergström & Chown, 1999; Gillespie & Roderick, 2002).

Collembola

Fossils of the earliest known hexapod and proposed ancestor of Collembola, *Rhyniella praecursor,* were discovered in Scotland and dated to the Devonian (380 Myr ago; Whalley & Jarzembowski, 1981). This early date suggests that Collembola was one of the first hexapod groups to evolve. Rapoport (1971) speculates that all the main families of Collembola were present before the fragmentation of Pangea and Gondwanaland. Furthermore, the recent phylogenetic placement of this group outside modern insects and in closer relation to the Crustacea makes it a crucial taxon in the study of the evolution of hexapods (Nardi *et al*, 2003).

With approximately 6500 described species and numerous undescribed taxa, Collembola is one of the most successful invertebrate orders (Hopkin, 1997). Members of this group are capable of surviving extreme temperatures allowing them to colonise a variety of environments including Antarctica and sub-Antarctic islands (Hopkin, 1997). Several morphological characters unite the group, such as the furca - a "spring" on the fourth abdominal segment that allows Collembola to be propelled forward at great distances in a very short space of time. Another important feature is the ventral tube which is not only crucial in maintaining fluid balances but also helps Collembola attach to surfaces (Hopkin, 1997).

Collembola are generally regarded as sedentary creatures because of their inability to fly. However, although their only methods of locomotion are walking and jumping, it has been shown that they can disperse over long distances via birds, human activities and especially by wind where they have been recorded at heights in excess of 3000m (Freeman, 1952 in Hopkin, 1997). Additionally, they can also survive extended periods of submersion in

seawater (Coulson *et al*, 2002) allowing them to colonise new islands by clinging to sea-surface debris carried on ocean currents.

Collembola feed on fungal hyphae or decaying plant material and with the exception of one species, are beneficial to the environment by controlling and influencing the growth of mycorrhizae (Hopkin, 1997). Gisin (1943) divided the springtails into three groups based on their general ecology. These groups are the euedaphic springtails (permanently underground), the hemiedaphic species (in the superficial soil layers and leaf litter) and the epiedaphic species (surface-living and on vegetation). Peterson (1980) found that the food consumed by the surface-living species tended to be more scattered but of higher quality than that of soil-living species, and perhaps the higher mobility of surface-living species allows them to find nutrient-rich food sources. This would suggest a more uncertain life-strategy for the epiedaphic Collembola when compared to the euedaphic species in their more constant and sheltered environment (Peterson, 2002). Peterson (1980) concluded that the more variable conditions at the surface would necessitate a high degree of genetic variation.

Collembola on Marion Island

Marion Island supports 22 species of vascular plants and 165 species of lichens, mosses and liverworts (Gremmen, 1981). These are grouped into 41 plant communities and futher grouped into six community complexes (Gremmen, 1981); Collembola occur in all of these (Gabriel *et al*, 2001). The Marion Island Collembola fauna is composed of seven families representing twelve genera and sixteen species (see Table 1.1).

Of the sixteen Collembola species that occur on Marion Island, ten are indigenous and four of these are endemic to the island. All three species chosen for this study belong to the Order Arthropleona. These include two indigenous species (*Cryptopygus antarcticus* and *Tullbergia bisetosa*) and a third species (*Isotomurus* cf. *palustris*), which is believed to have been introduced to the island by humans at least 25 years ago (Déharveng, 1981).

Table 1.1: Collembola fauna that occur on Marion Island (adapted from Déharveng, 1981). Species included in this study are in bold text. An asterisk (*) indicates introduced species.

Family	**Species**	**Distribution**
Neanuridae	*Friesea viennei* (Déharveng, 1981)	Sub-Antarctic
Hypogastruridae	*Ceratophysella denticulata* (Bagnall, 1941)*	Cosmopolitan
	Hypogastrura viatica (Tullberg, 1872)	Cosmopolitan
Onychiuridae	***Tullbergia bisetosa* (Börner, 1903)**	**Sub-Antarctic**
Isotomidae	*Isotoma* (Folsotoma) *marionensis* (Déharveng, 1981)	Endemic
	Cryptopygus dubius (Déharveng, 1981)	Endemic
	***Cryptopygus antarcticus travei* (Déharveng, 1981)**	**Endemic**
	Cryptopygus ceacus (Wahlgren, 1906)	Sub-Antarctic
	Cryptopygus tricuspis (Enderlein, 1909)	Sub-Antarctic
	Isotoma (Parisotoma) *notabilis* (Schäffer, 1896)*	Cosmopolitan
	Isotomurus palustris* (Müller, 1876)	**Cosmopolitan**
Tomoceridae	*Pogonognathellus flavescens* (Tullberg, 1871)*	Cosmopolitan
Neelidae	*Megalothorax* sp*	Unknown
Sminthuridae	*Sminthurinus tuberculates* (Delamare & Deboutteville & Massoud, 1963) = *Sminthurinus kerguelensis* (Salmon, 1964 sensu *S.* cf. *kerguelensis* (Déharveng, 1981)	Sub-Antarctic
	S*minthurinus granulosis* (Déharveng, 1981)	Sub-Antarctic
	Katianna marionensis n. sp.	Endemic

Cryptopygus antarcticus travei (Willem, 1901)

The genus *Cryptopygus* comprises 71 species (according to Hopkin, 1997). It belongs to the family Isotomidae, one of the most widespread and frequently recorded springtail families (Hopkin, 1997). The distribution of *C. antarcticus* is circumpolar (Broady, 1979). The subspecies *C. antarcticus travei* is endemic to Marion Island with the remaining four subspecies found on Antarctica, Crozet, Kerguelen and Heard Islands (Déharveng, 1981). Contrary to the findings of Gabriel *et al* (2001), *C. antarcticus* also occurs intermittently in plant communities other than the high-altitude mire (personal observation). This species feeds on a wide array of plant material including unicellular green algae and dead moss remains, as well as fungal hyphae (Block, 1985). The species' ability to feed on a range of readily accessible material is thought to contribute to its success in colonising a wide range of habitats (Broady, 1979).

Tullbergia bisetosa (Börner, 1903)

Tullbergia bisetosa is one of 19 species described for this genus (according to Hopkin, 1997) and belongs to the family Onychiuridae, superfamily Poduroidea. This group is commonly known as the blind Collembola and consists of many leaf-litter or soil-dwelling forms in which pigment, eyes and furca are almost always absent (Hopkin, 1997). While present on Marion Island, *T. bisetosa* also occurs on Kerguelen and Heard Islands (Déharveng, 1981).

Isotomurus cf. *palustris* (Müller, 1876)

Isotomurus (Family Isotomidae; Hopkin, 1997) is one of the most widespread Collembolan genera with a cosmopolitan distribution (Carapelli *et al*, 2004). The genus comprises 48 species (according to Hopkin, 1997), of which a large number occur in moist habitats on the surface such as litter layers (Carapelli *et al*, 1995; 1997).

The taxonomy of Collembola appears to be complicated and the position of Collembola relative to other arthropod groups is a topic of heated discussion (Hopkin, 1997). Among the more problematic groups within Collembola, the genus *Isotomurus* has been extensively studied (Carapelli *et al*, 1995; 1997; 2004). Within this genus alone there are numerous taxonomic problems. It is a widespread genus with an estimated number of 48 described species and subspecies based on diverse morphological characters (Carapelli *et al*, 1995; 2004).

Of the six *Isotomurus* species included by Carapelli (1995), *I. palustris* was characterized by the highest levels of polymorphism. In contrast to the other species that preferred certain microhabitats across the sampling site, *I. palustris* had a wider habitat selection leading to the hypothesis that it might have an elevated ability to adapt to various habitats.

Isotomurus palustris comprises many subspecies with various colour patterns (Carapelli *et al*, 2004). *Isotomurus.* cf. *palustris* is believed to have been introduced to Marion Island by human transportation of animal fodder at least 25 years ago (Déharveng, 1981). On Marion Island, the species occurs in all the lowland plant communities, the mid-altitude mires and is also occasionally found in fellfield habitat in the cushion plant *Azorella selago* (Gabriel *et al*, 2001) where it is the only introduced Collembola species (Barendse & Chown, 2001). *I.* cf *palustris* has also been introduced to Kerguelen Island (Déharveng, 1981).

Climate change and its effects on organisms

During the past 700 000 years, global temperature cycles have been characterised by cold glacial periods interspersed with short, warm interglacials. These changes in temperature have had different effects on organisms depending on factors such as distance from the equator, influence of the ocean and its currents, land mass size and location of the land and its mountain ranges (Hewitt, 1996). Interpretations of present genetic structures will therefore require consideration of the interaction between biology and geography as well as historical climatic shifts (Hewitt, 2000; Wiens & Donoghue, 2004).

Under the scenario of changes in available habitat, a species may potentially follow one of three routes. It could go extinct, disperse to more suitable habitat or, survive in small refugia from where it may expand again when conditions become more suitable (Hewitt, 2000). The topography of a region will play a major role in the postglacial spread of organisms. During glacial maxima, cold-intolerant taxa would be confined to refugia where suitable habitat remains. During warmer interglacials, these taxa would expand their range as new habitats become available and allow the coexistence of previously separated close relatives in the same habitat (Knowles, 2000). It is therefore not surprising that two of the most pressing issues to conservation are species' range expansions / contractions and the invasion of exotic organisms caused by climate changes (Bergström & Chown, 1999; Travis & Dytham, 2002).

The Southern Ocean plays an important role in global climate change where studies (Bergström & Chown, 1999; Smith, 1987; 2002) have shown that there has been a definite warming of the surface of the Southern Oceans during the last 50 years due to global warming (Gille, 2002). The few sub-Antarctic islands scattered throughout the Southern Ocean may be good indicators of historic glacial conditions caused by the fluctuation of the Antarctic Polar Frontal Zone (APFZ) (Hall, 1990) and may therefore provide ideal study areas to further understand the effect that climatic oscillations have had on population genetic structure.

The levels of genetic variation present on a single island in the sub-Antarctic may be very different from that observed for continental species because, at least from a geomorphological

perspective, the sub-Antarctic peri-glacial environment is very different from the seasonal frost and permafrost environments (Hall, 1990; Hall, 2002; Boelhouwers *et al*, 2002). This may have important consequences for the genetic population structure of organisms that occur on these islands.

Antarctic Collembola *Gressitticantha terranova* (Fanciulli *et al*, 2001), *Isotoma klovstadi* (Frati *et al*, 2001) and *Gomphiocephalus hodgsoni* (Stevens & Hogg, 2003) show higher levels of genetic differentiation on the continent compared to lower levels described on the island habitats surrounding Antarctica (Stevens & Hogg, 2003). It is believed that the higher levels of variation on the continent are caused by the fact that ideal living conditions for these animals are not uniformly distributed. Where conditions are thus favourable, large numbers of Collembola aggregate (Fanciulli *et al*, 2001) and are separated by inhospitable ice-sheets resulting in large genetic differences between conspecific populations.

There are two possible non-exclusive explanations for the lower levels of genetic variation found for species inhabiting islands surrounding Antarctica: first, these islands were probably only recently colonised (after the last glacial maximum) by a few individuals. Second, repeated bottleneck events may have caused ancient, variable island populations to lose more variation than the populations on mainland Antarctica particularly if the islands were colonised by relatively few individuals. These already lower levels of variation would have been even more reduced by repeated bottlenecks (Stevens & Hogg, 2003).

This high genetic divergence among Collembola populations on Antarctica concurs with the hypothesis that Collembola seem to have reduced dispersal capabilities (Frati *et al*, 2001). Similar high levels of genetic divergence were also found in European (Carapelli *et al*, 1997; Fanciulli *et al*, 2000; Van der Wurff *et al*, 2003) and Australian Collembola (Garrick *et al*, 2004). On Marion Island, nothing is known about the population genetic structure of Collembola. In this respect a population genetic study of selected Collembola would allow for the examination of population genetic structure in relation to habitat availability.

Phylogeography

Phylogeography is the study of relationships between, and distributions of alleles within a species (Avise *et al*, 1987) where an attempt is made to describe and discuss spatial patterns of genetic variation against a geographical background (Avise, 1987; Knowles & Maddison, 2002; Templeton; 2004). The major evolutionary forces responsible for shaping the current genetic structure of any species include mutation, migration, selection and drift (Cavalli-Sforza & Edwards, 1967; Hey & Nielsen, 2004). Additionally, the population structure and ecology of a species have an intricate interaction (Slatkin, 1985; King *et al,* 2003) where the current genetic structure of a taxon is also greatly influenced by past ecological processes. It follows that the current genetic structure would similarly influence both the present and future ecology of the population. First, lineages that are able to survive under certain conditions are favoured and second, genetic barriers between populations may become so large that they lead to speciation. Therefore, understanding the factors that influence the evolution of taxa in a certain region allows for predictions about the patterns of genetic diversity, which in turn, provide insight into the response of the population to current or future events.

Many key ecological and evolutionary characteristics of populations are greatly affected by the movement of individuals and genes through space (Slatkin, 1985; Hanski & Gilpin, 1997). Although theory gives us a very good idea of what could hypothetically happen with interruptions in gene flow, i.e. adaptation to certain habitats and eventually speciation, the actual amount of gene flow itself is very difficult to measure in reality (Slatkin, 1987; Hey & Nielsen, 2004).

One of the first population genetic hypotheses was that of "isolation by distance" whereby the further apart two populations are geographically; the greater should be their expected genetic differentiation (Wright, 1943; Pogson *et al*, 2001). There are two most prominent models that were developed to explain genetic variation in populations: the island model -where "…the variance of gene frequencies among different populations should be related to the number of migrants which come into each population each generation" (Wright, 1931; Whitlock & McCauley, 1999; Stevens & Hogg, 2003), and the stepping stone model where "…the population is divided into colonies and the migration of each generation is restricted to nearby colonies" (Kimura & Weiss, 1964; Wang *et al*, 2004). In his development of the island model, Sewall Wright introduced F-statistics as a handy way of summarising population structure where F_{IT} summarises the difference between the individual and the total population, F_{ST} the difference between each subpopulation and the total population and F_{IS}, the difference between the individual and the subpopulation (Wright, 1922; 1951).

Since the introduction of F_{ST}, many parameters have been derived from it such as G_{ST} (Vandewoestijne *et al*, 1999), and Φ_{ST} (Tzeng *et al*, 2004) with the same assumptions as F_{ST} (Whitlock & McCauley, 1999). These assumptions are firstly that there is an infinite number of populations with N diploid individuals, and that each of these populations gives and receives a fraction *m* of its individuals each generation to and from the migrant pool. Secondly, that all populations are equally likely to give and receive migrants from all other populations. Thirdly, that there is no selection or mutation and lastly, that each population persists indefinitely.

Most natural populations do in fact violate some if not all of these assumptions. This makes it crucial to understand the underlying biology and history of the species in order to make inferences from the data. For example, although the amounts of gene flow may be similar in two different populations, it may have different effects on each population where drift and selection play different roles. When the geographic range of a species stays the same and local populations continue for a long time, gene flow occurs mostly through the movement of individuals amongst established populations and will thus prevent genetic differentiation. This process is also critically dependant on population size and in a situation where populations are unstable (e.g. due to frequent extinction and recolonisation or changes in geographical range), gene flow and population subdivision will have an important evolutionary role and could lead to significant population genetic structure (Slatkin, 1987). Also, statistically significant differences do not always indicate biologically significant differences and vice versa (Hedrick, 1999).

Haplotype networks give information on both the geographical distribution of alleles as well as their evolution through time. Templeton (1998) describes a unique method of using this information to determine population structure using nested clade analysis. Although this method is limited, it is a way to determine whether a certain geographical grouping is statiscally significant (Templeton,2004). A summary of the rational behind the method can be read in Templton (1998).

In order for scientists to be able to make inferences about the current molecular population structure, a molecular clock can be applied to data in order to put a date to the observed divergence. Usually, species with a rich fossil record can be used to calibrate the molecular clock and the clock can then be applied to estimate divergence times for similar species that don't have a rich fossil record. For Collembola an exact calibration does not exist, but it has been estimated that arthropod mtDNA evolves at between 2% and 2.3% per million years for COI specifically and the entire mitochondrial genome in general (Soto-Adames, 2002; Stevens & Hogg, 2003; Ballard & Whitlock, 2004). In the absence of fossil calibration points,

molecular clock methods represent the only way of estimating times of divergences. Although not ideal, it does allow rough estimates of divergence times so that some conclusions as to the reason for divergence can be drawn. Nevertheless, molecular clocks have been severely critiqued (Alroy, 1999; Lee, 1999; Conroy & Van Tuinen, 2003; Hedges & Kumar, 2004; Near *et al*, 2005). In addition, because there is such a large error connected to time estimates derived from DNA sequence divergence (Hewitt, 1996), low divergence times among populations bear high standard deviations often in excess of the divergence period itself (Knowles & Maddison, 2002).

Molecular markers

Appropriate markers should reveal similarity within taxonomic units as well as differences among them. There are many properties that the ideal marker should have (Cruickshank, 2002). However, there is no single gene that will have all these properties. Certain regions on the mitochondrial genome exhibit considerable levels of variation allowing for the documentation of differences within and among populations (Parker *et al*, 1998; Ballard & Whitlock, 2004). Another advantage to using mitochondrial DNA at the population level is its maternal mode of inheritance (single-copy and therefore likely to have a fast rate of evolution) which allows us to trace genealogies through time (Ballard & Whitlock, 2004). It is therefore often the marker of choice and has been used extensively for detecting molecular differences both at species (Roehrdanz *et al*, 2002) and population levels (Ballard & Whitlock, 2004). Examples of phylogeographic studies based on mtDNA include humans (Fuselli *et al*, 2003), other mammals (Seddon *et al*, 2001; Lloyd, 2003; Li *et al*, 2003; Ruedi & Castella, 2003; Alpers *et al*, 2004), reptiles (Creer *et al*, 2001; Lesia *et al*, 2003; Berry & Gleeson, 2004), frogs (Nielson *et al*, 2001), fish (Viñas *et al*, 2004; Wang *et al*, 2004), insects (Hale & Sing, 1991; Brown *et al*, 1996; Vandewoestijne *et al*, 2004), Collembola (Frati *et al*, 2000; Fanciulli *et al*, 2000; Garrick, 2002; Garrick *et al,* 2004; Stevens & Hogg, 2003) as well as other invertebrates (Blouin *et al*, 1995; Turner *et al*, 2000; Tzeng *et al*, 2004; Tolley *et al,* 2005).

Among insects, the evolution of mitochondrial DNA has been extensively studied (Wolstenholme & Clary, 1985; Liu & Beckenbach, 1992; Frati *et al*, 1997; Simmons & Weller, 2001) and the complete mitochondrial sequences for several species are now available for further studies and comparisons (Crozier & Crozier, 1993; Nardi *et al*, 2001; Ballard & Whitlock, 2004). Cytochrome b and the control region are more frequently used in vertebrate studies, but cytochrome oxidase I (COI) and cytochrome oxidase II (COII) seem to be more often favoured in studies on insects (Miura *et al*, 2000; Simmons & Weller, 2001; Vandergast *et al*, 2004; Vandewoestijne *et al*, 2004). COI and COII form subunits of the cytochrome oxidase c complex where oxidative phosphorylation takes place (Campbell, 1999), whereas

the other subunits of this complex are encoded by the nuclear genome (Campbell, 1999). As these gene products are extremely important, their genes are very well conserved among all metazoan taxa (Hale & Singh, 1991; Ballard & Whitlock, 2004), although the order of the genes varies among different taxa of invertebrates (Crozier & Crozier, 1993; Nardi *et al,* 2001*;* Roehrdanz *et al*, 2002). A study by Vandewoestijne and colleagues (2004) also found that, in butterflies, the control region displays less genetic variability than COI.

In this study, sequences from COI and COII were used to determine the population structure of Collembola on Marion Island. Notwithstanding reports about the conserved nature of COI (Simon *et al*, 1994; Vandewoestijne *et al*, 2004), it has been successfully used to determine the genetic structure of various Collembola species from New Zealand (Garrick, 2002; 2004) and Antarctica (Stevens & Hogg, 2003). In addition Frati *et al* (1997) found substantial genetic divergences between families and species of Collembola based on COII gene sequences and speculated that either COII is evolving intrinsically faster in Collembola than in other insect orders or that it simply reflects the age of the species and families.

Chapter 2

Materials and Methods

Materials and Methods

Sample Collection

Collembola specimens included in the present study were collected on Marion Island during the relief voyages in April of 2003 and April 2004. Mosses and other plant material were sampled from localities across the island (see Table 2.1 and Figure 2.1). Specimens were collected using mechanical aspirators (hand made) and preserved in absolute ethanol. Analyses were conducted in the Evolutionary Genomics Laboratory, Department of Botany and Zoology, Stellenbosch University.

Table 2.1: Sampling localities included in the present study. Geographic coordinates as well as sampling altitude are indicated. Locality numbers (indicated in brackets) correspond with those used in the figures. The number of specimens collected from each sampling locality are indicated in the last three columns.

Locality Name	Longitude	Latitude	Altitude	*C. antarcticus*	*T. bisetosa*	*I. palustris*
Trypot (1)	46˚53'05''S	37˚52'05''E	13m	0	0	8
Skua Ridge (2)	46˚52'04''S	37˚50'17''E	88m	0	0	4
Swartkops (3)	46˚55'28''S	37˚35'44''E	57m	9	5	9
Kildalkey Bay (4)	46˚58'01''S	37˚31'10''E	19m	11	5	9
Blue Petrel Bay (5)	46˚50'48''S	37˚49'06''E	33m	0	5	11
Ship's Cove (6)	46˚51'14''S	37˚50'30''E	30m	6	0	0
Fred's Hill (7)	46˚54' 49''S	37˚50'21''E	230m	0	0	5
Hendrik Vister (8)	46˚53'12''S	37˚48'49''E	282m	11	5	0
Katedraal Krans (9)	46˚53'54''S	37˚46'29''E	768m	5	3	0
Long Ridge (10)	46˚52'55"S	37˚47'11"E	515m	5	0	0
Archway Bay (11)	46˚53'56''S	37˚53'42''E	39m	10	2	8
Greyheaded (12)	46˚57'43''S	37˚42'31''E	84m	7	0	0
Mixed Pickle (13)	46˚52'20''S	37˚38'21''E	50m	0	4	5
Cape Davis (14)	46˚49'41''S	37˚42'14''E	63m	7	5	0
Rook's Bay (15)	46˚58'01''S	37˚39'39''E	67m	0	3	0
Stony Ridge (16)	46˚55'03"S	37˚51'31"E	162m	7	0	0
Mesrug (17)	46˚56'37"S	37˚50'52"E	141m	0	3	0
Bullard (18)	46˚55'16''S	37˚52'53''E	42m	0	0	6

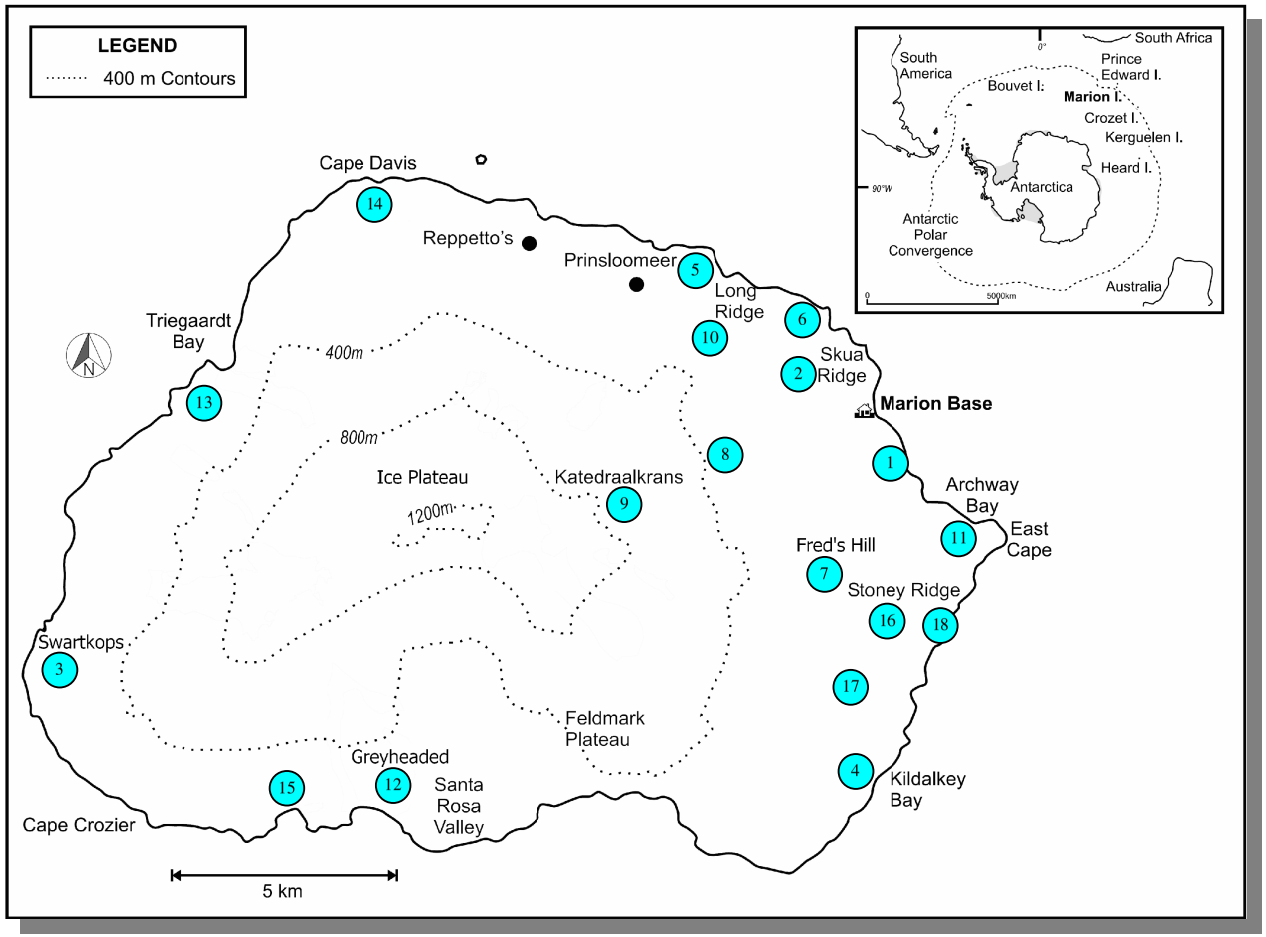

Figure 2.1: Map of Marion Island redrawn from Nel *et al*, 2003. Circles indicate the sampling localities across Marion Island, for all species included in this study. Locality numbers correspond to those given in Table 2.1.

DNA Extraction

Total genomic DNA was extracted using standard procedures described in Maniatis *et al* (1982). Whole animals were digested in 550µl of DNA lysis buffer (100mM NaCl, 10mMTris-Cl, 1mM EDTA), 30% SDS solution and 20µl proteinase K (10mg/ml). The mixture was incubated at 55ºC for four hours, where upon the DNA was extracted using standard phenol/chloroform: isopropanol protocols (Hillis *et al*, 1996). Dried DNA pellets were resuspended in 30µl of 0.5 M TE and stored at –4°C until required.

DNA Amplification and Sequencing

Standard polymerase chain reactions (PCR) were set up to a total volume of 25µl in the presence of 3µl of 25mM $MgCl_2$ 1x buffer, 0.05mM dNTP solution, 0.1µM forward and reverse primers, 1 unit of *Taq* polymerase (Super Therm, Southern Cross Biotechnology) and 1-2µl of extracted DNA. For the COI gene amplification, the primer combination HCO2198 and LCO1490 was used (Folmer *et al*, 1994) while COIIa and COIIb (Frati *et al*, 1997) were used for COII amplifications. Primer sequences are provided in Table 2.2. Amplification of the COII gene for *T. bisetosa* required the synthesis of species-specific primers and the sequences for these are also shown in Table 2.2. The temperature regime for all PCR reactions was: 94ºC for 4 min; (94ºC for 40s; 42ºC for 30s; 72ºC for 45s) for 36 cycles. A final annealing and extension cycle (42ºC for 5 min followed by 72ºC for 10 min) completed the reactions.

All samples were checked for successful PCR amplification by electrophoreses in 1% TAE agarose gels containing ethidium bromide and visualised under a UV light. COI-gene amplicons were directly cycle-sequenced using PCR product, 2µl (quarter reaction) of Big Dye terminator chemistry (version 3; Perkin Elmer, Applied Biosystems) and 0.5µl of a 10µM solution of each primer (HCO2198 and LCO1490) respectively. COII amplicons contained non-specific amplification (multiple bands). For this gene fragment total PCR products were electrophoresed and the correct COII amplification band (determined through sequencing and BLAST searches in GenBank) was cut from the gel and purified with the Wizard® SV Gel and PCR Clean-Up System (Promega). Sequencing followed the same protocol as described for the COI gene. Unincorporated dye label was removed by sephadex columns before the samples were run on an ABI 3100 automated sequencer (Applied Biosystems). Sequences were checked using Sequence Navigator (ABI, version 1.01) and aligned by eye. Sequences were submitted to GENBANK under accession numbers DQ147289-DQ147558.

Table 2.2: Primer sequences targeting both the COI and COII genes employed in this study. Species-specific primers were developed for *T. bisetosa* and are indicated with an asterisk (*)**.**

Primer Name	Gene	Sequence
HCO 2198	COI	5' TAAACTTCAGGGTGACCAAAAAATCA 3'
LCO 1490	COI	5' GGTCAACAAATCATAAAGATATTGG 3'
COIIa	COII	5' AATATGGCAGATTAGTGCA 3'
COIIb	COII	5' GTTTAAGAGACCAGTACTT 3'
T-COIIa*	COII	5' ATGAAAATTACTTCCTTTAG 3'
T-COIIb*	COII	5' GCACTGGCCAAAGTATAAGC 3'

Data Analyses

Nucleotide composition and genetic distances were determined using PAUP* (Swofford, 2000). The number of variable and parsimony-informative sites as well as haplotype and nucleotide diversity was determined using DnaSP version 3 (Rozas & Rozas, 1999).

Traditionally, evolutionary relationships are graphically presented as phylogenetic trees. These evolutionary reconstructions are very appealing as they create a visual image, but are strongly dependent on the evolutionary model that was applied in their estimation (Cavalli-Sforza & Edwards, 1967; Gaggiotti *et al*, 2004). Trees are further created under the assumption that all the branches are bifurcating (Tajima, 1983); this is not strictly true when conspecific populations are under study because ancestral haplotypes do not go extinct when new ones arise. To account for these assumptions, and to investigate the level of connectedness among populations, we constructed a minimum spanning network by hand using distances (mutational steps separating haplotypes) obtained from Arlequin version 2 (Schneider *et al*, 2000).

Additionally, a more demographic approach was adopted where the spatial distribution of mitochondrial variation was explored using an analysis of molecular variance (AMOVA; Excoffier *et al*, 1992 as implemented in Arlequin version 2; Schneider *et al*, 2000). Permutational procedures (1000 randomisations) were used to provide significance tests for the F_{ST} statistics, as well as for each of the hierarchical variance components. A hierarchical AMOVA was performed for each of the three species where localities were grouped according to natural geographic barriers that are present across the island. This was done to test whether these human-perceived barriers affect gene flow in Collembola. A SAMOVA (spatial analysis of molecular variance; Dupanloup *et al*, 2002) analysis was also performed to

determine the strongest population-genetic structure and as a by-product points to possible geographic barriers to gene flow between these inferred groups.

Possible deviations in allele frequency from equilibrium were investigated with Fu's F test statistic (Fu, 1997; Arlequin version 2). This test, which incorporates information from the haplotype distribution, is relatively sensitive to frequency changes yet insensitive to small population sizes (Ramos-Onsins & Rozas, 2002). Deviations from equilibrium may be caused by mutations, migrations and / or selection. A significant Fs value might indicate possible changes in historical population demography. To further investigate the historical demography of Collembola populations on Marion Island, a mismatch analysis was performed (Rogers & Harpending, 1992) (i.e., the distribution of observed pairwise nucleotide site differences against their respective frequency of occurrence) as implemented in DnaSP version 3 (Rozas & Rozas, 1999). Population expansions and contractions are characterised by signatures in the patterns of molecular diversity (Harpending *et al,* 1998; Schneider & Excoffier, 1999). A rapid change in population size will produce an even distribution of pairwise differences resulting in a Poisson distribution. In contrast, stable populations are characterised by multimodal distributions. A $\chi 2$ goodness-of-fit test (implemented in Arlequin version 2) was used to compare observed and expected distributions.

A nested clade analysis (NCA) was performed to test for statistically significant phylogeographical associations. For this, only connections between haplotypes that had a probability greater than 95% were considered (as calculated in Parsprob v. 1.1; Posada, 1998). The minimum spanning network was converted into a nested clade design using standard nesting rules (Templeton *et al*, 1987; Templeton & Sing, 1993). Once the minimum spanning network has been converted, each clade is nominated by C-N, (C is the nesting level of the clade and N is the number of a particular clade at a given nesting level) and then the geographical data are computed directly from the longitude and latitude records. Two distance statistics are given, namely Dc (clade distance) and Dn (nested clade distance). Dc is the geographical range of a given clade whereas Dn is the average distance that an individual with a certain haplotype lies from the centre of all the individuals from the same clade (Templeton, 1998). Differences in distance measures between tip clades and clades immediately interior to them are significant in shaping the geographical structure of genetic variation (Templeton *et al*, 1995). Nested-contingency analysis and nested geographical distance analyses were implemented with GEODIS 2.0 (Posada *et al,* 2000) using 1000 permutations. The results were interpreted for patterns of population structure and historical events using the latest version of Templeton's (2004) inference key available at: *http://zoology.byu.edu/crandall_lab/dposada/documents/NCA-key(24Oct01).pdf* .

A test for isolation-by-distance was performed by implementing the Mantel test in Arlequin version 2 (Schneider *et al*, 2000). Longitudinal and latitudinal coordinates were converted into geographical distances and compared to the observed genetic distances between populations. Because it was reasoned that the animals would cross the island by circumventing the ice plateau, rather than traversing it, distances were not always calculated as straight-line measurements, but rather followed the most likely colonisation routes available to the animals. The correlation-coefficient values as calculated by Arlequin (Version 2; Schneider *et al*, 2000) were taken as genetic differences.

Chapter 3

Results

Results

Results for each of the three species will be presented separately. All analyses were performed for the gene (COI and COII) fragments singly as well as combined. Unfortunately the COII gene failed to amplify for *T. bisetosa* despite exhaustive amplification attempts. This included amplification under various temperature regimes and $MgCl_2$ titrations as well as designing several sets of primers from aligned Collembola species (including *I. palustris*) available in Genbank (further details are provided below). Our analyses for *T. bisetosa* are therefore based only on the COI gene.

3.1 *Cryptopygus antarcticus*

3.1.1 COI gene

A 610 base pair fragment was amplified and sequenced for COI. Of these 610 characters ten (1.64%) were polymorphic, five (0.82%) of which were parsimony informative. The ratios of A:C:G:T calculated in PAUP*(Swofford, 2000) were 0.34: 0.20: 0.21 and 0.25, respectively.

For each of the 10 sampling localities included for *C. antarcticus* (Figure 3.1), five or more specimens were sequenced giving a total of 75 specimens in our analyses of the COI gene. Thirty-three different haplotypes were found for the 75 individuals analysed. Although a large number of private / unique haplotypes were observed (27 specimens had a unique haplotype), several specimens shared haplotypes and this is reflected in the haplotype diversity of 0.69. The most common haplotype was represented in 41.33% of the sequenced individuals and was widespread across the island.

The highest uncorrected sequence divergence of 3.04% between *C. antarcticus* specimens recorded for the COI gene was between specimens sampled at Greyheaded (12) and Archway Bay (11). The average uncorrected pairwise nucleotide divergence values separating localities are shown in Table 3.1. The highest nucleotide divergence (average) between any of the sampling localities was 1.32% ± 0.43% between Archway Bay (11) and Kildalkey Bay (4). Nucleotide diversity for the whole island, calculated in DnaSP version 3 (Rozas & Rozas, 1999), was 0.00754.

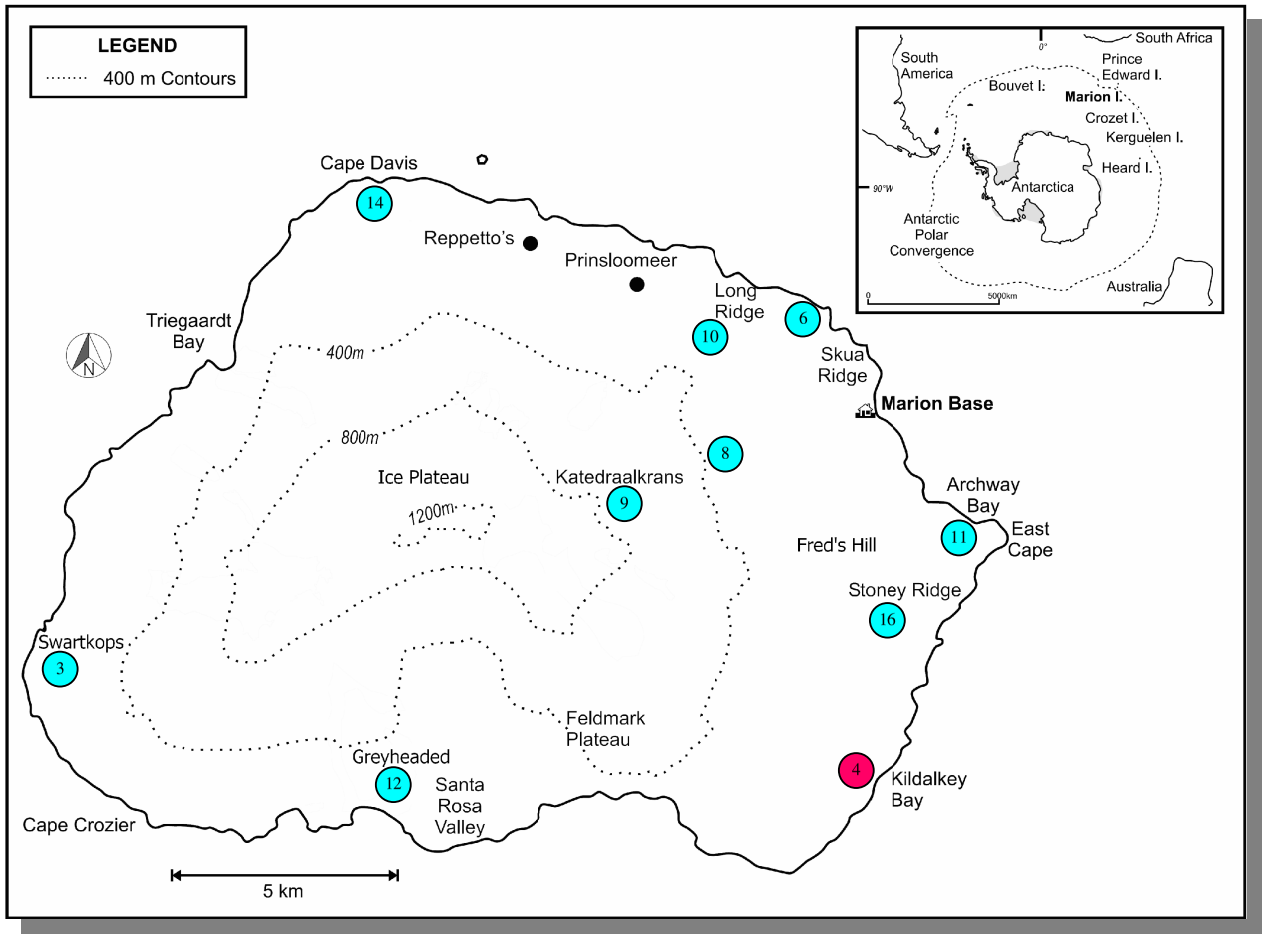

Figure 3.1: Map of Marion Island redrawn from Nel *et al*, 2003. Circles indicate the sampling localities for *C. antarcticus* specimens. Locality numbers correspond to those given in Table 2.1. Kildalkey Bay(4) is shown in a different colour to highlight the difference between it and the other localities.

Table 3.1: Average percentage sequence divergence separating localities for *C. antarcticus* are shown below the diagonal. These values are based on 610 bp of the COI gene. Average uncorrected divergence between specimens within localities is indicated on the diagonal.

	Swartkops	**Kildalkey**	**Hendrik Vister**	**Katedraal**	**Cape Davis**	**Ship's Cove**	**Archway Bay**	**Greyheaded**	**Long Ridge**	**Stony Ridge**
Swartkops	0.22 ±0.20%									
Kildalkey	0.97 ±0.25%	0.66 ±0.33%								
Hendrik Vister	0.41 ±0.39%	0.98 ±0.29%	0.57 ±0.46%							
Katedraal	0.70 ±0.36%	0.73 ±0.45%	0.74 ±0.37%	0.80 ±0.36%						
Cape Davis	0.21 ±0.20%	1.03 ±0.30%	0.41 ±0.38%	0.76 ±0.45%	0.18 ±0.11%					
Ship's Cove	0.68 ±0.41%	0.83 ±0.28%	0.74 ±0.42%	0.71 ±0.26%	0.70 ±0.42%	0.68 ±0.47%				
Archway Bay	0.40 ±0.32%	1.32 ±0.43%	0.64 ±0.54%	0.95 ±0.49%	0.38 ±0.28%	0.99 ±0.55%	0.55 ±0.35%			
Greyheaded	0.55 ±0.47%	0.80 ±0.37%	0.64 ±0.46%	0.67 ±0.38%	0.56 ±0.48%	0.71 ±0.37%	0.83 ±0.66%	0.69 ±0.46%		
Long Ridge	0.26 ±0.17%	1.14 ±0.30%	0.49 ±0.43%	0.83 ±0.42%	0.25 ±0.16%	0.84 ±0.49%	0.46 ±0.30%	0.64 ±0.55%	0.31 ±0.13%	
Stony Ridge	0.22 ±0.19%	1.09 ±0.31%	0.46 ±0.43%	0.80 ±0.43%	0.23 ±0.18%	0.80 ±0.48%	0.41 ±0.29%	0.63 ±0.54%	0.26 ±0.16%	0.26 ±0.18%

Fu's Fs-value applied to the population across the entire island, an indication of whether the species is in Hardy-Weinberg equilibrium, was calculated in Arlequin (version 2; Schneider *et al*, 2000). This test returned a highly significant value (–26.04; $p<0.001$) indicating a deviation from equilibrium. To investigate whether this deviation may be the result of a demographic event such as a population expansion, a mismatch distribution was performed in DnaSP version 3 (Rozas & Rozas, 1999). There was significant ($p<0.001$) concurrence (Arlequin, version 2; Schneider *et al*, 2000) between the mismatch distribution data (see Figure 3.2) and the expected values under a model of sudden expansion and this model (of sudden expansion) could therefore not be rejected.

The results of the AMOVA analysis (Arlequin, version 2; Schneider *et al*, 2000) are shown in Table 3.2. When considering sampling localities as populations (no groups were defined), the largest portion of the variance (71%) is attributed to variation within different populations with the remainder (29%) ascribed to variation between populations. Despite high variability within populations, the F_{ST} value of 0.290 was significant with a p-value of less than 0.001. For the spatial analyses of molecular variation (SAMOVA; Dupanloup *et al*, 2002) it is noteworthy that F_{CT} values were decidedly significant ($p<0.001$) and Kildalkey Bay (4) was shown to be significantly different from the rest of the localities. The only exception was when two groups were specified for this analysis, the F_{CT} value (0.39) was high, but not significant ($p<0.09$). The F_{CT} value increases and remains significant as higher numbers of groups are specified, up to seven, where after the values either decrease or become insignificant. Kildalkey Bay (4) always remained distinct irrespective of how many groups the localities were divided into.

The minimum spanning network, drawn using the least number of mutational steps between haplotypes calculated in Arlequin (version 2; Schneider *et al*, 2000) is shown in Figure 3.3. Specimens are identified by two numbers; the first is the locality number and the second the specimen number from each respective sampling locality. Two distinct demographic signatures are evident: the first resembles a star-like pattern typical of a recent expansion with the majority of specimens (32) sharing a haplotype. Several haplotypes, separated by only one or two nucleotide differences, are derived from this central haplotype (following the nesting rules by Templeton *et al*, 1995). The second signature, indicative of a more ancient expansion, is characterised by higher numbers of nucleotide differences separating haplotypes. In this second pattern, fewer specimens share haplotypes.

A Mantel test, as implemented in Arlequin (version 2; Schneider *et al*, 2000), was done to detect possible isolation by distance. Distances calculated (see materials and methods) for the most likely colonisation routes followed by Collembola on the island are presented in Table

3.3. The Mantel test yielded negative, non-significant correlation values indicating no isolation by distance for *C. antarcticus* based on the COI data set.

For the nested clade analysis (Templeton *et al*, 1995) only one of the clades (Clade 3-1, Figure 3.3) had a significant p-value ($p<0.05$) showing that there is a possibility of geographic partitioning, but none of the Dc nor Dn distances within the nested clade were significantly large or significantly small; the outcome of this analysis was inconclusive. The sequence of the inference key can be seen in Table 3.4.

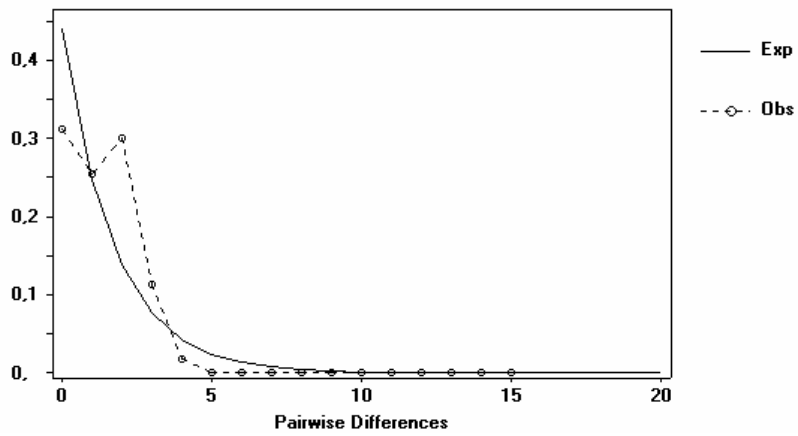

Figure 3.2: Mismatch distribution indicating the number of differences separating haplotypes plotted against their respective frequencies. The solid line indicates the expected values and the broken line the obesreved values. This analysis is based on 610 bp of the COI gene for *C. antarcticus*.

Table 3.2: AMOVA results based on 610 bp of COI for 75 *C. antarcticus* specimens where each sampling locality has been analysed as a separate population.

Source of variation	**df**	**Sum of Squares**	**Variance Components**	**Percentage of Variation**
Among Populations	9	55.58	0.62 Va	28.94
Within Populations	65	99.71	1.53 Vb	71.06
Total	74	155.29	2.16	
Fixation index	F_{ST}: 0.29	($p<0.001$)		

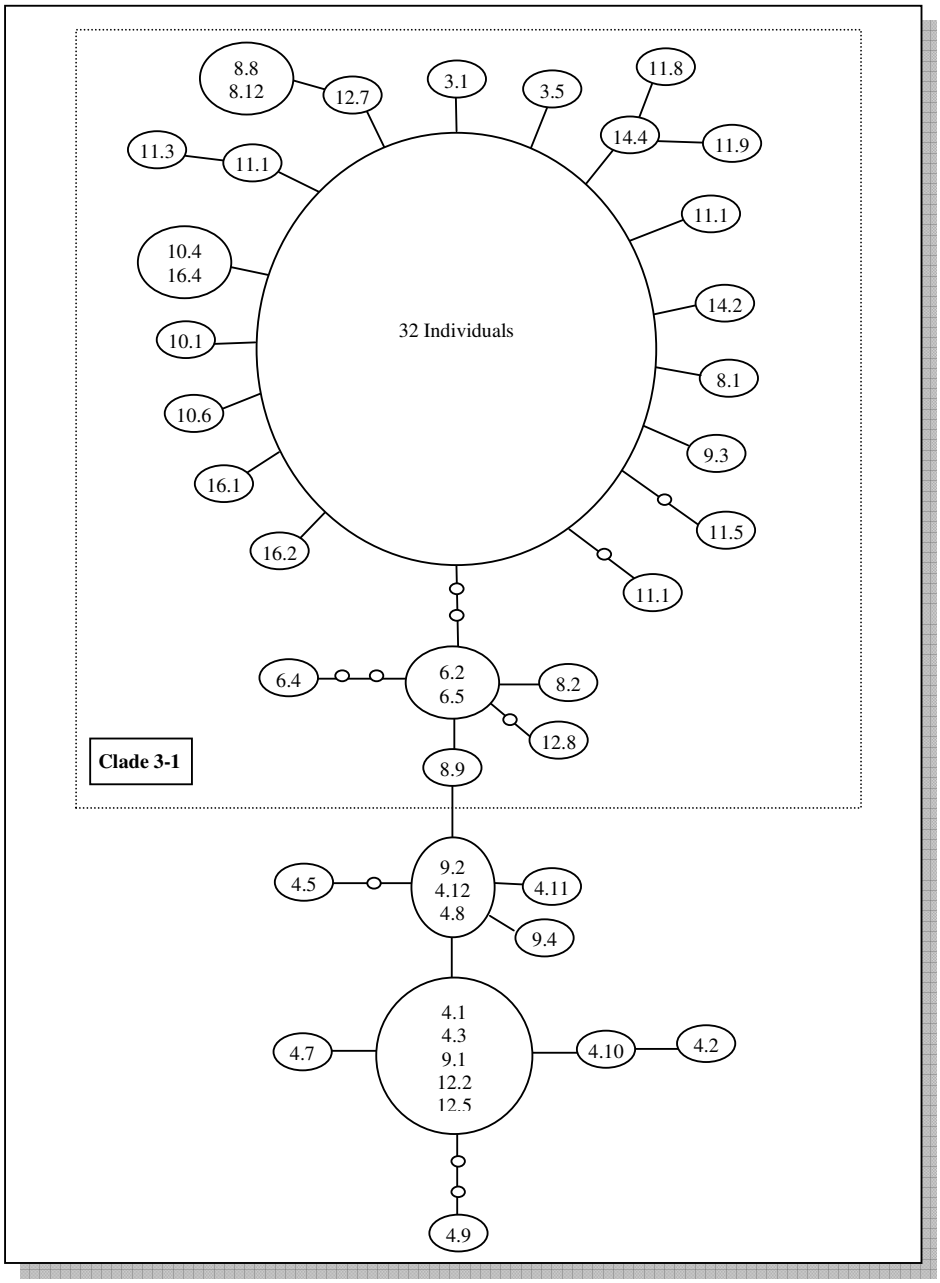

Figure 3.3: The minimum spanning network based on 610 bp of COI for 75 *C. antarcticus* specimens. Lines represent one mutational difference between haplotypes. Missing haplotypes are indicated by circles. Circles representing haplotypes are drawn according to scale with specimens characterised by haplotypes within (except where a large number of specimens shared a haplotype in which case only the number of specimens is indicated). (The dotted box shows the clade 3-1 that showed significant geographic partitioning in the nested cladistic analysis).

Table 3.3: Geographical distances (given in kilometres) separating sampling localities for *C. antarcticus*. Distances are calculated along the most likely colonisation routes.

	Swartkops	**Kildalkey Bay**	**Hendrik Vister**	**Katedraal**	**Cape Davis**	**Ship's Cove**	**Archway Bay**	**Greyheaded**	**Long Ridge**	**Stony Ridge**
Swartkops										
Kildalkey Bay	20.9									
Hendrik Vister	22.5	9								
Katedraal	17	3.4	3.4							
Cape Davis	13	7.8	9.6	7.7						
Ship's Cove	23.9	11.9	3.9	7.1	11					
Archway Bay	28.9	8.2	6.4	9.6	21.9	6.4				
Greyheaded	9	11.7	11.4	7.9	14.9	15.1	15.7			
Long Ridge	20.3	10.6	2.1	2.5	12.2	5	8.5	10.6		
Stony Ridge	5.3	5.3	3.9	5.9	13.5	6.4	4.3	11.4	5.9	

Table 3.4: Inference chain based on results of geographical dispersion analysis. *http://zoology.byu.edu/crandall_lab/dposada/documents/NCA-key(24Oct01).pdf* was used as source for inference key. Only those clades that resulted in a rejection of the null hypothesis of panmixia are included in this table.

Clade	**Chain of inference**	**Inference**
Haplotypes nested in 3-1	1-2-11-17-NO	Inconclusive outcome

3.1.2 COII gene

For the COII gene, a fragment of 620 base pairs was analysed. Of these sequenced characters, 18 were variable sites and 7 were parsimony-informative. Base frequencies as calculated in PAUP* (Swofford, 2000) were A:C:G:T = 0.30: 0.21: 0.15 and 0.34 respectively.

For the COII gene, 48 individuals were sequenced from the same 10 localities as described for the COI gene (see Figure 3.1). These 48 specimens, represent a subset (apart from four additional individuals) of those included in the COI analyses, were characterized by 22 haplotypes. Of these 22 haplotypes, 17 (77%) were private / unique haplotypes with a haplotype diversity of 0.79. The most common haplotype was found in 21 specimens (43.75%); this haplotype is widespread across the island and is found in the majority of localities sampled. Nucleotide diversity for the species a whole, calculated in DnaSP version 3 (Rozas & Rozas, 1999), was 0.00387.

The highest average sequence divergence between specimens for the COII gene was 1.22% ± 0.53% between two specimens both sampled at Katedraal (9). The highest average pairwise nucleotide divergence between two different localities was 0.87 ± 0.42% between Katedraal (9) and Kildalkey Bay (4) (Table 3.5). In the analysis of COI for the same species, Kildalkey Bay (4) also showed the highest average pairwise difference between two localities.

An Fs-value of –26.61 ($p<0.001$) was calculated in Arlequin version 2 (Schneider *et al*, 2000) for the species as a whole and this highly significant value shows a departure from equilibrium possibly signifying a demographic change. The mismatch-distribution (DnaSP version 3; Rozas & Rozas, 1999) confirms a recent demographic change by showing a good fit to the expected values under a population expansion model (Figure 3.4).

Results of the AMOVA (calculated in Arlequin version 2; Schneider *et al*, 2000) are shown in Table 3.6. When considering sampling localities as populations, 12% of the variation can be attributed to variation among populations whereas the rest of the variation (88%) lies within populations. Although only a small fraction of the variation is accounted for by differences between populations, $F_{ST} = 0.12$ was significant ($p<0.001$). In the spatial analysis of molecular variance (SAMOVA, Dupanloup *et al*, 2002), Kildalkey Bay (4) again stood out as being significantly different from the rest of the localities with the F_{CT} value calculated at 0.20 ($p<0.001$). When sampling localities are divided into two or more groups: Kildalkey Bay (4) groups by itself and the rest of the localities fall in groups two and three.

Table 3.5: Average percentage of sequence divergence separating localities for *C. antarcticus* shown below the diagonal. These values are based on 620 bp of the COII gene. Average uncorrected divergence between specimens within localities is indicated on the diagonal.

	Swartkops	**Kildalkey**	**Hendrik Vister**	**Katedraal**	**Cape Davis**	**Ship's Cove**	**Archway Bay**	**Greyheaded**	**Long Ridge**	**Stony Ridge**
Swartkops	0.55 ±0.10%									
Kildalkey	0.62 ±0.20%	0.32 ±0.17%								
Hendrik Vister	0.46 ±0.26%	0.42 ±0.20%	0.33 ±0.28%							
Katedraal	0.73 ±0.24%	0.87 ±0.42%	0.68 ±0.49%	1.22 ±0.53%						
Cape Davis	0.43 ±0.18%	0.41 ±0.19%	0.33 ±0.25%	0.69 ±0.41%	0.25 ±0.17%					
Ship's Cove	0.58 ±0.30%	0.51 ±0.33%	0.48 ±0.36%	0.83 ±0.44%	0.51 ±0.32%	0.57 ±0.39%				
Archway Bay	0.41 ±0.16%	0.39 ±16%	0.28 ±0.23%	0.68 ±0.45%	0.29 ±0.14%	0.41 ±0.25%	0.22 ±0.11%			
Greyheaded	0.49 ±0.19%	0.43 ±0.22%	0.34 ±0.23%	0.73 ±0.44%	0.39 ±0.19%	0.45 ±0.29%	0.30 ±0.15%	0.38 ±0.18%		
Long Ridge	0.42 ±0.24%	0.40 ±0.24%	0.35 ±0.30%	0.79 ±0.54%	0.32 ±0.23%	0.52 ±0.35%	0.28 ±0.20%	0.39 ±0.22%	0.38 ±0%	
Stony Ridge	0.42 ±0.27%	0.40 ±0.28%	0.35 ±0.34%	0.67 ±0.49%	0.29 ±0.25%	0.44 ±0.30%	0.29 ±0.26%	0.36 ±0.25%	0.32 ±0.31%	0.29 ±0.26%

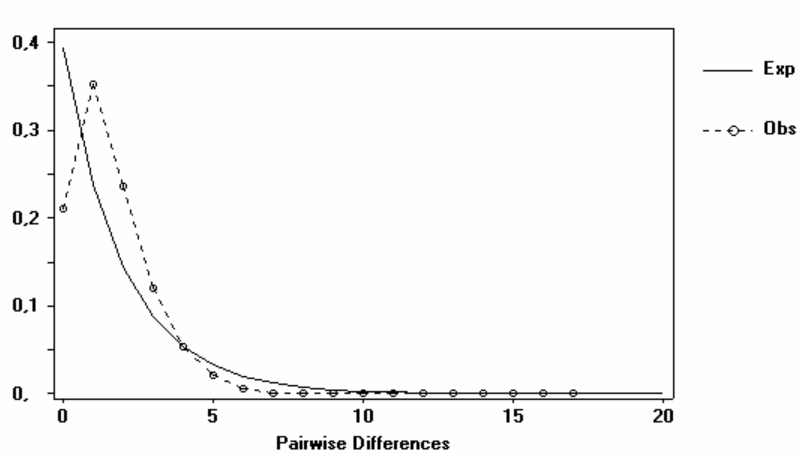

Figure 3.4: Mismatch distribution plotting the number of differences separating haplotypes against their respective frequencies. The solid line indicates the expected values and the broken line the obesreved values. Analysis based on 620 bp of the COII gene for *C. antarcticus*.

Table 3.6: AMOVA results based on 620 bp of the COII gene for 48 *C. antarcticus* specimens. Sampling localities were considered as different populations.

Source of variation	**df**	**Sum of Squares**	**Variance Components**	**Percentage of Variation**
Among Populations	9	18.16	0.16 Va	11.74
Within Populations	38	47.04	1.24 Vb	88.26
Total	47	65.20	1.40	
Fixation index	F_{ST}: 0.12	(p=<0.001)		

The minimum spanning network, depicting the least number of changes between haplotypes (calculated in Arlequin version 2; Schneider *et al*, 2000) is shown in Figure 3.5. There is a central haplotype shared by 22 individuals with the derivatives radiating from the central haplotype connecting by no more than four mutational steps.

The Mantel test (Arlequin version 2; Schneider *et al*, 2000), using geographical distances calculated along the most likely colonisation routes (Table 3.3), gave a non-significant correlation value thus indicating no isolation by distance for this gene in *C. antarcticus.*

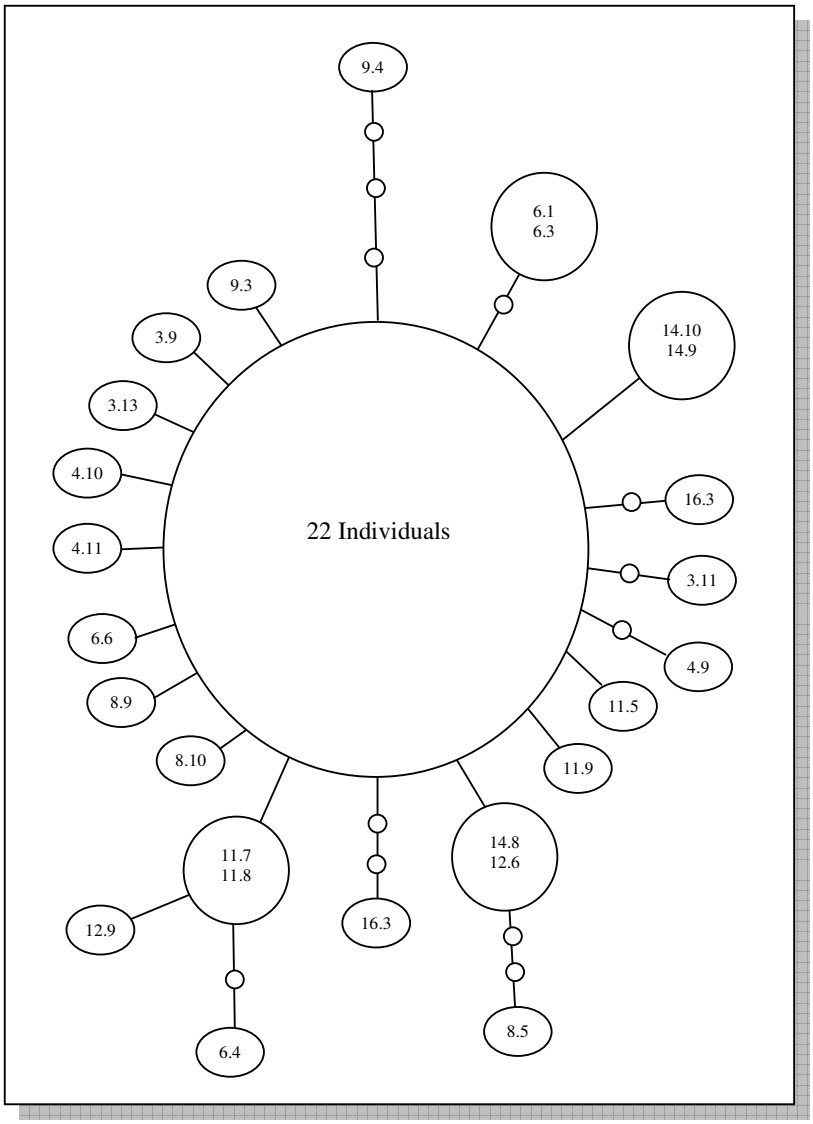

Figure 3.5: Minimum spanning network based on 620 bp of COII for 48 *C. antarcticus* specimens. Lines represent one mutational difference between haplotypes. Missing haplotypes are indicated by circles. Circles representing haplotypes are drawn according to scale with specimens characterised by haplotypes within (except where a large number of specimens shared a haplotype in which case only the number of specimens is indicated).

3.1.3 Combined data set

Forty-four individuals for which there were comparable data available for both the COI and COII genes were combined to give a final combined data set of 1160 base pairs because non-informative characters were cut off from the beginning of the COII gene and the end of the COI gene. The combined data set is characterized by 28 (2.4%) variable and 9 (0.77%) parsimony-informative sites and the base frequencies (PAUP*; Swofford, 2000) were: A:C:G:T = 0.32: 0.20: 0.18 and 0.29 respectively.

Forty different haplotypes characterized the 44 specimens included for the combined data set. Five haplotypes were shared by 2 individuals respectively with the remaining 34 haplotypes being unique to single specimens. Strikingly, for the combined data set, almost all specimens displayed a unique haplotype as reflected in the haplotype diversity of 0.96.

The highest sequence divergence (1.09 ± 0.50%) separating individuals based on the combined data set was between specimens occurring at a single locality (Katedraal Krans (9)); these were the same two individuals that were found to have the highest sequence divergence based on the COII data set. The largest average pairwise nucleotide difference (0.91 ± 0.32%) between two different localities was found between Katedraal (9) and Long Ridge (10) (Table 3.7). Nucleotide diversity (calculated in DnaSP version 3) was 0.00503.

Cryptopygus antarcticus across Marion Island deviated from Hardy-Weinberg equilibrium as indicated by Fu's F-statistic (–25.222; $p<0.001$; Arlequin version 2, Schneider *et al*, 2000) for the species across the island. When differences between haplotypes are plotted against respective frequencies (mismatch analysis; DnaSP version 3, Rozas & Rozas, 1999) there is a good correspondence between observed values and those expected under a demographic change ($p<0.001$; Arlequin version 2, Schneider *et al*, 2000).

The results for the AMOVA (Arlequin version 2, Schneider *et al*, 2000) are shown in Table 3.8. When considering sampling localities as populations, 78% of the variation lies within populations ($F_{ST} = 0.22$; $p<0.001$). Similar to the findings from the separate gene fragments, a combined data spatial analysis of variance (SAMOVA; Dupanloup *et al*, 2002) confirmed the uniqueness of Kildalkey (4). Variation among localities ($F_{CT} = 0.30$; $p<0.001$) was maximized under a seven-group scenario where $F_{CT} = 0.33$ ($p<0.02$). In the three-group scenario, F_{CT} is only marginally lower at 0.30 ($p<0.02$) and even in this configuration Kildalkey Bay (4) grouped on its own. It is interesting to note that Katedraal (9) also groups on its own when four groups are specified.

Table 3.7: Average percentage of sequence divergence separating localities for *C. antarcticus* shown below the diagonal. These values are based on 1160 bp of the combined data. Average uncorrected percentage sequence divergence between specimens within localities is indicated on the diagonal.

	Swartkops	**Kildalkey**	**Hendrik Vister**	**Katedraal**	**Cape Davis**	**Ship's Cove**	**Archway Bay**	**Greyheaded**	**Long Ridge**	**Stony Ridge**
Swartkops	0.29 ±0.07%									
Kildalkey	0.82 ±0.15%	0.42 ±0.22%								
Hendrik Vister	0.42 ±0.26%	0.74 ±0.20%	0.50 ±0.33%							
Katedraal	0.75 ±0.33%	0.79 ±0.40%	0.77 ±0.36%	1.09 ±0.50%						
Cape Davis	0.27 ±0.10%	0.79 ±0.13%	0.40 ±0.26%	0.80 ±0.28%	0.23 ±0.10%					
Ship's Cove	0.61 ±0.30%	0.66 ±0.29%	0.65 ±0.32%	0.79 ±0.31%	0.63 ±0.30%	0.64 ±0.43%				
Archway Bay	0.42 ±0.12%	0.89 ±0.14%	0.49 ±0.27%	0.86 ±0.24%	0.34 ±0.10%	0.72 ±0.30%	0.41 ±0.12%			
Greyheaded	0.52 ±0.31%	0.62 ±0.26%	0.54 ±0.33%	0.77 ±0.31%	0.49 ±0.32%	0.62 ±0.29%	0.58 ±0.35%	0.63 ±0.32%		
Long Ridge	0.31 ±0.12%	0.87 ±0.14%	0.47 ±0.28%	0.91 ±0.32%	0.30 ±0.12%	0.70 ±0.34%	0.39 ±0.12%	0.57 ±0.35%	0.38 ±0%	
Stony Ridge	0.27 ±0.14%	0.81 ±0.19%	0.41 ±0.29%	0.81 ±0.34%	0.23 ±0.14%	0.60 ±0.32%	0.36 ±0.17%	0.50 ±0.33%	0.30 ±0.18%	0.27 ±0%

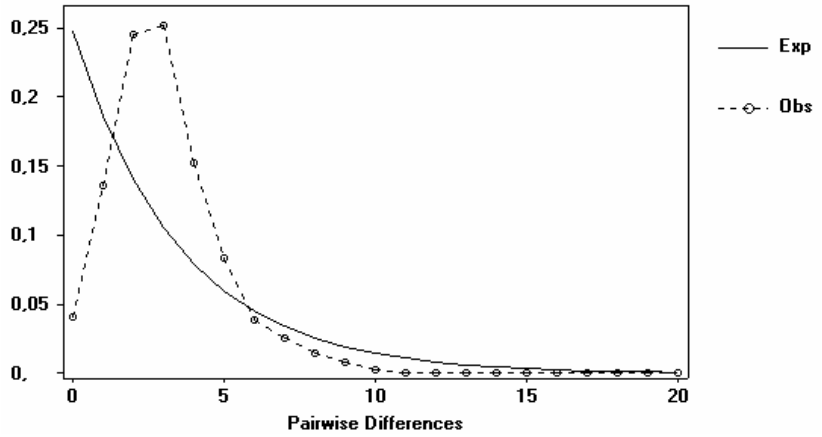

Figure 3.6: Mismatch distribution plotting the number of differences separating haplotypes against their respective frequencies. The solid line indicates the expected values and the broken line the obesreved values. Analysis is based on 1160 bp of the combined data for *C. antarcticus*.

Table 3.8: AMOVA results based on 1160 bp of the combined sequences for 44 *C. antarcticus* specimens. Sampling localities were considered as separate populations.

Source of variation	df	Sum of Squares	Variance Components	% Variation
Among Populations	9	55.40	0.77 Va	21.61
Within Populations	34	95.36	2.80 Vb	78.39
Total	43	150.76	3.58	
Fixation index	F_{ST}: 0.22	(p<0.001)		

When any number of groups higher than two and lower than nine was specified, the Fct values were significant and Kildalkey Bay always grouped on its own.

A minimum spanning network for the combined data set, based on the number of mutational steps separating haplotypes calculated in Arlequin (version 2, Schneider *et al*, 2000), is shown in Figure 3.7. The highest number of mutational steps separating haplotypes was six, and given that up to 14 mutational steps fall within the 95% parsimony probability (Parsprob v. 1.1; Posada, 1998), the network in its entirety was accepted. Two distinct demographic scenarios are evident from the network. The first, seen at the top of the network, is characterized by a central haplotype with many singletons connecting to it. This kind of pattern may represent a relatively recent expansion / colonisation / range extension (Kingman, 1980; Vandewoestijne *et al,* 2004). The second signature, seen in the bottom half of the cladogram, may represent an older and more stable pattern with up to six mutational events separating haplotypes.

Following the nesting rules for the nested clade analysis (Templeton *et al*, 1995), five of the one-step clades were geographically informative. NCA (Templeton *et al*, 1995) revealed significant geographical associations within only two clades; one at the two-step level (Clade 2-3) and one at the three-step level (Clade 3-2; Figure 3.7). The outcomes (steps) from following the inference key are shown in Table 3.9. The inferences made for Clade 2-3 lead to an inconclusive outcome, and the result for Clade 3-2 arrived at a result of contiguous range expansion.

The geographical distances (calculated along possible colonisation routes; Table 3.3) used to test for isolation by distance are the same as those used when the genes were analysed separately. The correlation values calculated in the Mantel test (Arlequin version 2, Schneider *et al*, 2000) gave a non-significant value showing that there is no correlation between the geographical and genetic distances.

Table 3.9: Inference chain based on results of geographical dispersion analysis. *http://zoology.byu.edu/crandall_lab/dposada/documents/NCA-key(24Oct01).pdf* was used as source for inference key. Only those clades that resulted in a rejection of the null hypothesis of panmixia are included in this table.

Clade	Chain of inference	Inference
Haplotypes nested in 2-3	1-2-11-17-NO	Inconclusive outcome
Haplotypes nested in 3-2	1-2-11-12-NO	Contiguous Range Expansion

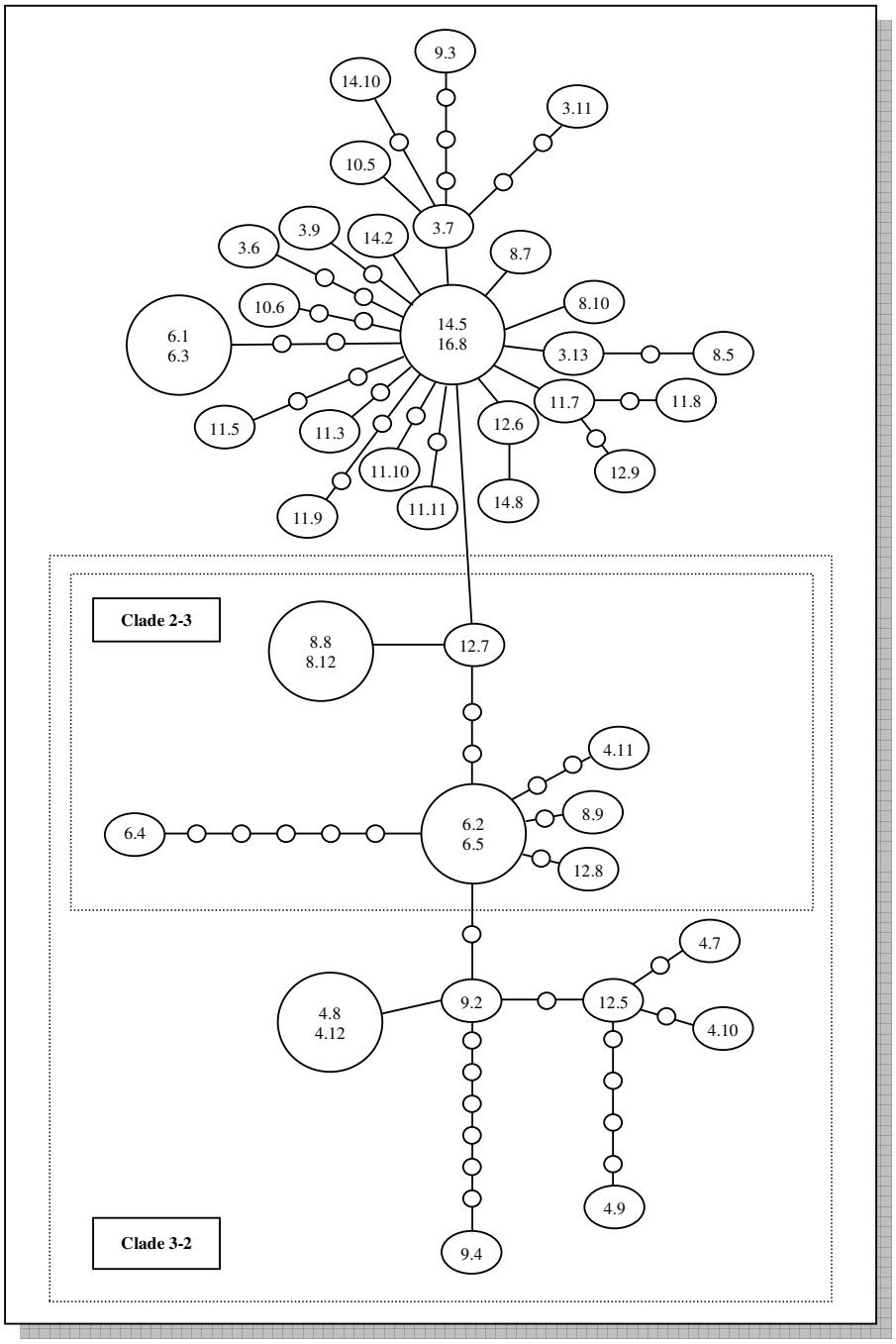

Figure 3.7: Minimum spanning network based on 1160 bp for 44 *C. antarcticus*. Lines represent one mutational difference between haplotypes. Missing haplotypes are indicated by circles. Circles representing haplotypes are drawn according to scale with specimens characterised by haplotypes within (except where a large number of specimens shared a haplotype in which case only the number of specimens is indicated). (The dotted boxes show the clades 3-1 and 2-3 that showed significant geographic partitioning in the nested cladistic analysis).

3.2 *Tullbergia bisetosa*

3.2.1 COI gene

Forty *T. bisetosa* specimens were sampled from ten localities across Marion Island (Figure 3.8). In total, 606 bp were analysed of which four were both variable and parsimony informative. Base frequencies calculated in PAUP* (Swofford, 2000) for A:C:G:T respectively were 0.33: 0.17: 0.22 and 0.28.

Thirteen haplotypes (32.5%) were identified for the 40 specimens analyzed which included nine singletons (22.5%). Haplotype diversity (DnaSp version 3, Rozas & Rozas, 1999) was 0.77. The three most common haplotypes characterized seven, seven and 15 individuals respectively; these three haplotypes were geographically widespread across Marion Island.

The highest percentage uncorrected nucleotide divergence between two specimens was 1.73% between specimens from Rook's Bay (15) and Mesrug (17). The average pairwise nucleotide divergence separating localities are shown in Table 3.10. The highest uncorrected average nucleotide divergence (0.83 ± 0.32%) was recorded between Rook's Hut (15) and Hendrik Vister (8) (see Figure 3.8 for localities). Nucleotide diversity, calculated in DnaSP version 3 (Rozas & Rozas, 1999) was 0.00848.

Arlequin version 2 (Schneider *et al*, 2000) determined an Fs-value of –8.728 ($p<0.001$), indicating a deviation from equilibrium for *T. bisetosa* for the species as a whole across the island. Results from the mismatch distribution (DnaSP version 3, Rozas & Rozas, 1999). Figure 3.9 shows that the observed population data fit reasonably well to that of expected values from a population under the model of sudden expansion. This would suggest that *T. bisetosa* was dynamic in the recent past.

AMOVA (Arlequin version 2, Schneider *et al*, 2000) results are shown in Table 3.11. Considering sampling localities as populations, the largest share of the variance (87%) is attributed to the within-population component with 13% ascribed to variation among populations. For the spatial analysis of molecular variance (SAMOVA; Dupanloup *et al*, 2002) for *T. bisetosa*, F_{CT} values increase as the number of groups rise from two to nine. When the localities are divided into two groups ($F_{CT} = 0.22$; $p<0.02$), Kildalkey Bay (4) and Rook's Hut (15) are grouped together separate from the remainder of the sampling localities. As the number of groups increases, Kildalkey Bay (4) and Rook's Hut (15) are always grouped together until the number of groups reaches seven, after which Kildalkey Bay (4) stands alone.

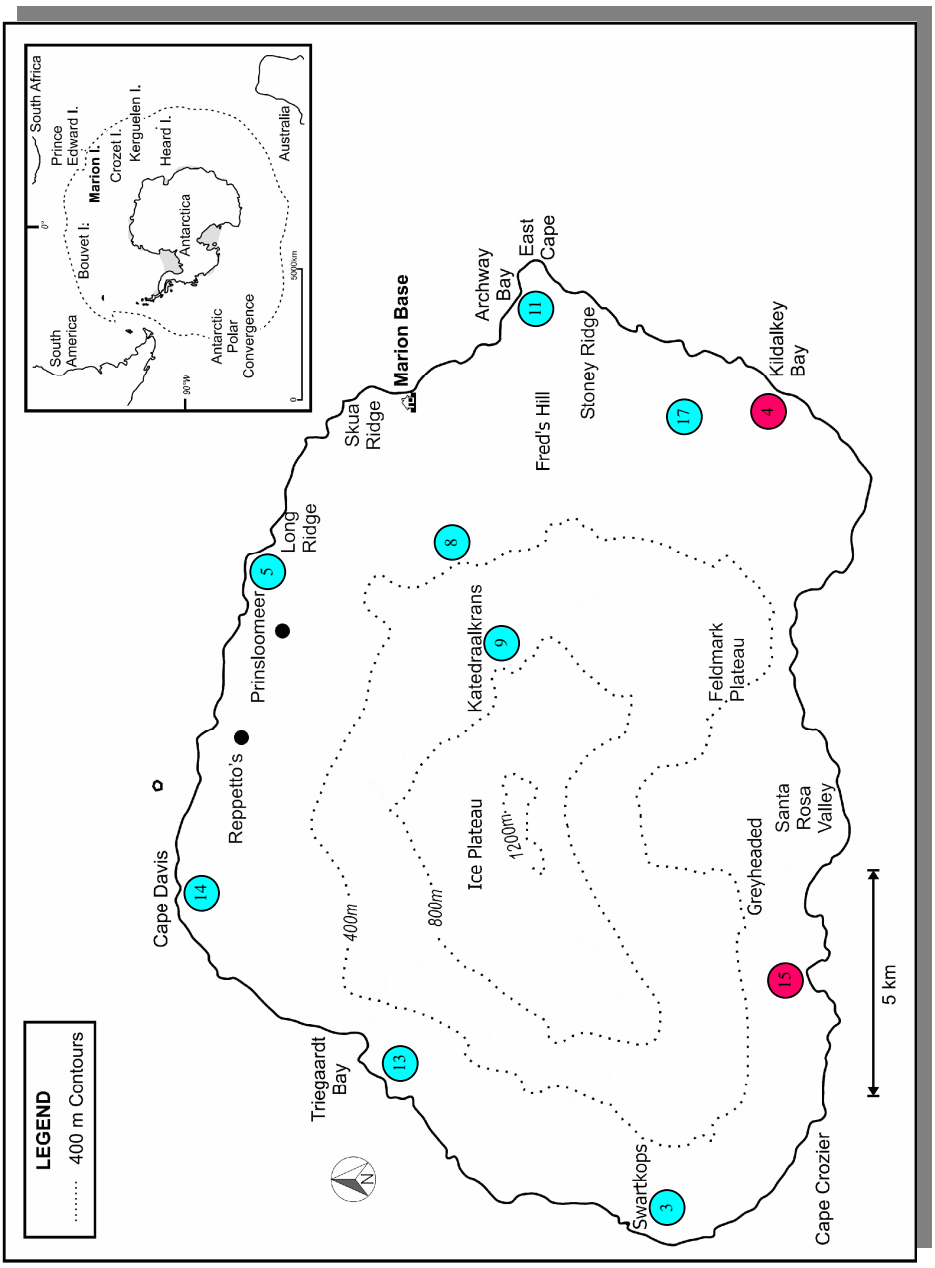

Figure 3.8: Map of Marion Island redrawn from Nel *et al*, 2003. Circles indicate the sampling localities for *T. bisetosa.* Locality numbers correspond to those given in Table 2.1. Kildalkey Bay (4) and Rook's hut (15) are shown in a different colour to highlight the differences among localities (see text for explanation).

Table 3.10: Average percentage of sequence divergence separating localities for *T. bisetosa* shown below the diagonal. These values are based on 606 bp of the COI gene. Average uncorrected divergence between specimens within localities are indicated on the diagonal.

	Swartkops	**Kildalkey**	**Hendrik Vister**	**Katedraal**	**Archway**	**Blue Petrel**	**Rook's Bay**	**Mixed Pickle**	**Cape Davis**	**Mesrug**
Swartkops	0.56±0.35%									
Kildalkey	0.54±0.29%	0.13±0.10%								
Hendrik Vister	0.45±0.35%	0.65±0.24%	0.31±0.20%							
Katedraal	0.45±0.37%	0.32±0.24%	0.58±0.39%	0.54±0.32%						
Archway	0.60±0.31%	0.43±0.22%	0.60±0.32%	0.48±0.21%	0.38±0%					
Blue Petrel	0.43±0.32%	0.44±0.27%	0.38±0.30%	0.44±0.33%	0.55±0.24%	0.42±0.27%				
Rook' s Bay	0.68±0.33%	0.32±0.14%	0.83±0.32%	0.44±0.33%	0.56±0.39%	0.63±0.30%	0.73±0.23%			
Mixed Pickle	0.50±0.37%	0.45±0.31%	0.46±0.38%	0.51±0.35%	0.55±0.25%	0.42±0.34%	0.61±0.43%	0.60±0.41%		
Cape Davis	0.62±0.30%	0.63±0.21%	0.54±0.33%	0.58±0.31%	0.59±0.26%	0.52±0.33%	0.74±0.34%	0.59±0.31%	0.63±0.25%	
Mesrug	0.72±0.30%	0.58±0.44%	0.67±0.23%	0.61±0.39%	0.54±0.30%	0.62±0.29%	0.73±0.54%	0.64±0.32%	0.70±0.27%	0.79±0.26%

Table 3.11: AMOVA results based on 606 bp of the COI gene for 40 *T. bisetosa* specimens. Sampling localities were considered as different populations.

Source of variation	df	Sum of Squares	Variance Components	Percentage of Variation
Among Populations	9	20.59	0.21Va	12.53
Within Populations	30	43.77	1.46 Vb	87.47
Total	39	64.36	1.67	
Fixation index	0.13	(p<0.05)		

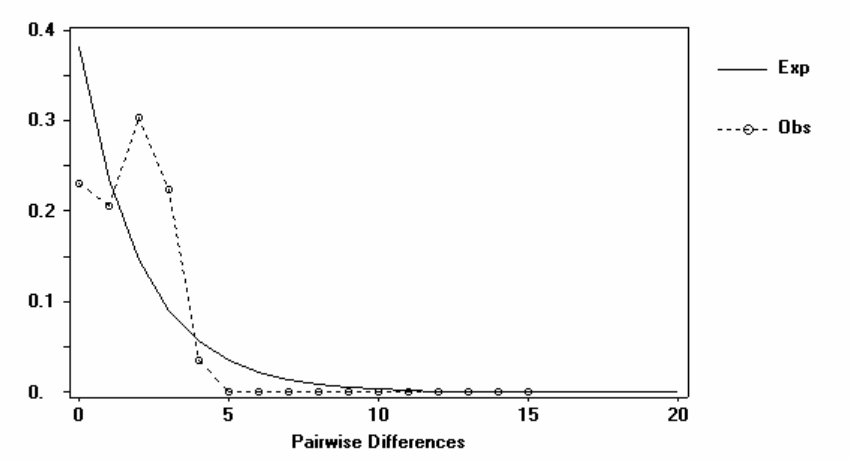

Figure 3.9: Mismatch distribution plotting the number of differences separating haplotypes against their respective frequencies. The solid line indicates the expected values and the broken line the obesreved values. Analysis is based on 606 bp of the COI gene for *T. bisetosa*.

The minimum spanning network for *T. bisetosa*, drawn by hand using the minimum number of mutational steps calculated by Arlequin (version 2, Schneider *et al*, 2000) is shown in Figure 3.10. The majority of specimens are characterized by three common haplotypes. These three common haplotypes are widely distributed across the geographical range of the species on Marion Island. Nine private alleles were detected for *T. bisetosa*. Haplotypes are separated by either one or two mutational steps.

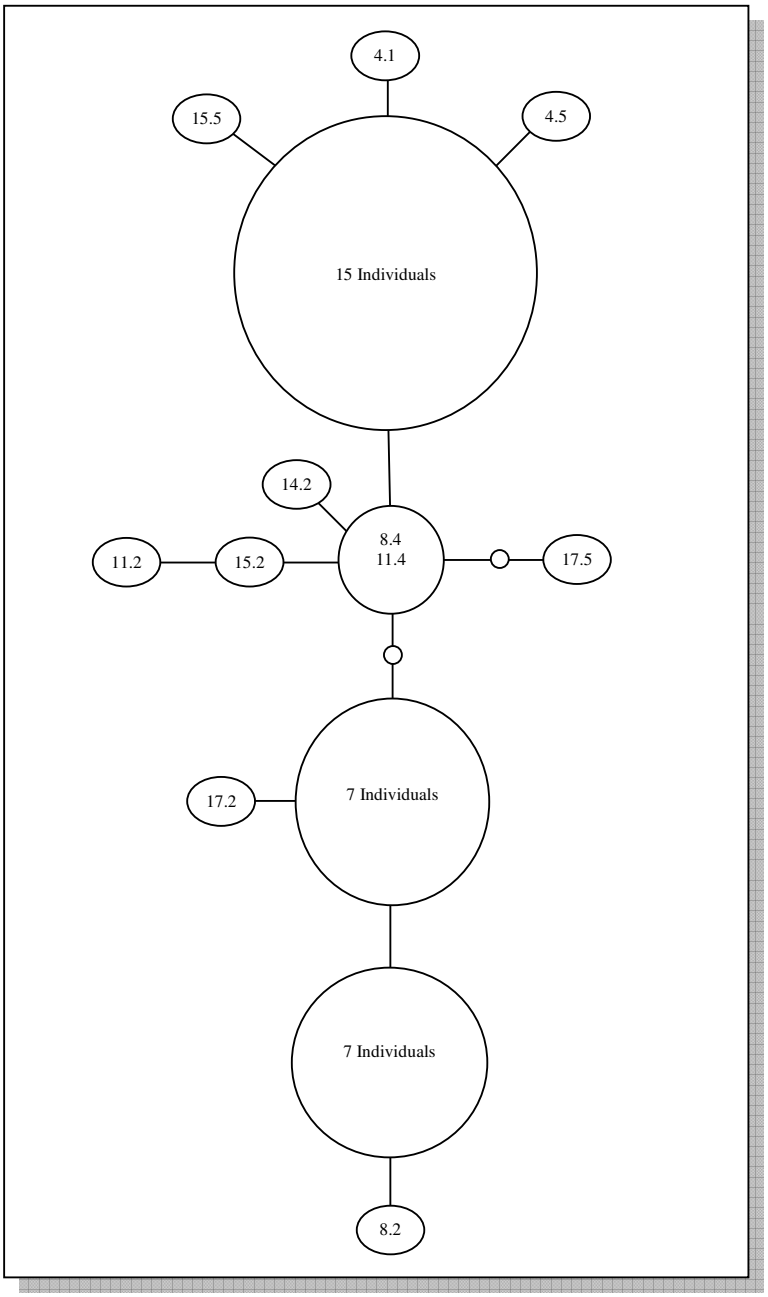

Figure 3.10: Minimum spanning network based on 606 bp of COI for 40 individuals of *T. bisetosa* specimens. Lines represent one mutational difference between haplotypes while missing haplotypes are indicated by circles. Circles representing haplotypes are drawn according to scale with specimens characterised by haplotypes within (except where a large number of specimens shared a haplotype in which case only the number of specimens is indicated).

Geographical distances separating *T. bisetosa* sampling localities are shown in Table 3.12. Similar to the results presented for *C. antarcticus*, these distances were calculated along the most likely colonisation routes rather than 'as the crow flies'. The Mantel test (Arlequin version 2, Schneider *et al*, 2000), to investigate possible isolation by distance, was not significant thus indicating no correlation between geographic and genetic distances.

The minimum spanning network was converted into nesting clades following Templeton *et al*, 1995 and Templeton, 1998. Nested clade analysis was performed and none of the clades revealed significant geographic partitioning.

Table 3.12: Geographical distances (given in kilometres) separating sampling localities for *T. bisetosa* calculated along the most likely colonisation routes.

	Swartkops	**Kildalkey Bay**	**Hendrik Vister**	**Katedraal**	**Archway Bay**	**Blue Petrel Bay**	**Rook's**	**Mixed Pickle**	**Cape Davis**	**Mesrug**
Swartkops										
Kildalkey Bay	20.9									
Hendrik Vister	22.5	9.0								
Katedraal	17.0	3.4	3.4							
Archway Bay	28.9	8.2	6.4	9.6						
Blue Petrel Bay	20.9	13.5	4.6	3.5	3.9					
Rook's	7.5	14.3	13.8	10.4	18.2	16.8				
Mixed Pickle	7.1	13.6	15.4	13.5	27.8	13.0	10.6			
Cape Davis	13.0	7.8	9.6	7.7	21.9	18.0	16.5	5.9		
Mesrug	19.7	3.7	6.0	6.7	5.9	10.1	13.3	20.3	14.5	

3.2.1 COII gene

Sequencing of the COII gene proved problematic for *T. bisetosa*. After it was found that COIIa and COIIb (Frati *et al*, 1997) failed to amplify the COII gene for *T. bisetosa*, a reverse primer was designed from gene sequences for Collembola species including *I. palustris* deposited in GENBANK. To produce a forward primer, LCO 1490 was reverse complemented and used in combination with the newly designed reverse primer. When that failed to work, a degenerate reverse primer was designed from the same sequences in GENBANK, but this also failed to amplify COII. Amplifications with general insect primers (COII-C2-N-3661 and COII-C2-J-3138; Simon *et al*, 1994) were also attempted, but without success. COII R-Lys and COII F-Leu (Mekawa *et al*, 1999; Miura *et al*, 2000) showed some amplification, but even after repeated cleaning up of PCR products, the sequencing reactions always showed multiple, indistinguishable bands. Sequences for the primers designed by us can be seen in Table 2.2.

3.3 *Isotomurus* cf. *palustris*

3.3.1 COI gene

For *I* .cf. *palustris*, a 480 bp fragment was sequenced for 65 specimens from nine localities across Marion Island (see Figure 3.11). The base frequencies for COI, calculated in PAUP* (Swofford, 2000) were as follows: A:C:G:T = 0.29: 0.19: 0.26 and 0.26 respectively. There were no polymorphisms in this region and only a single haplotype was found across the entire island.

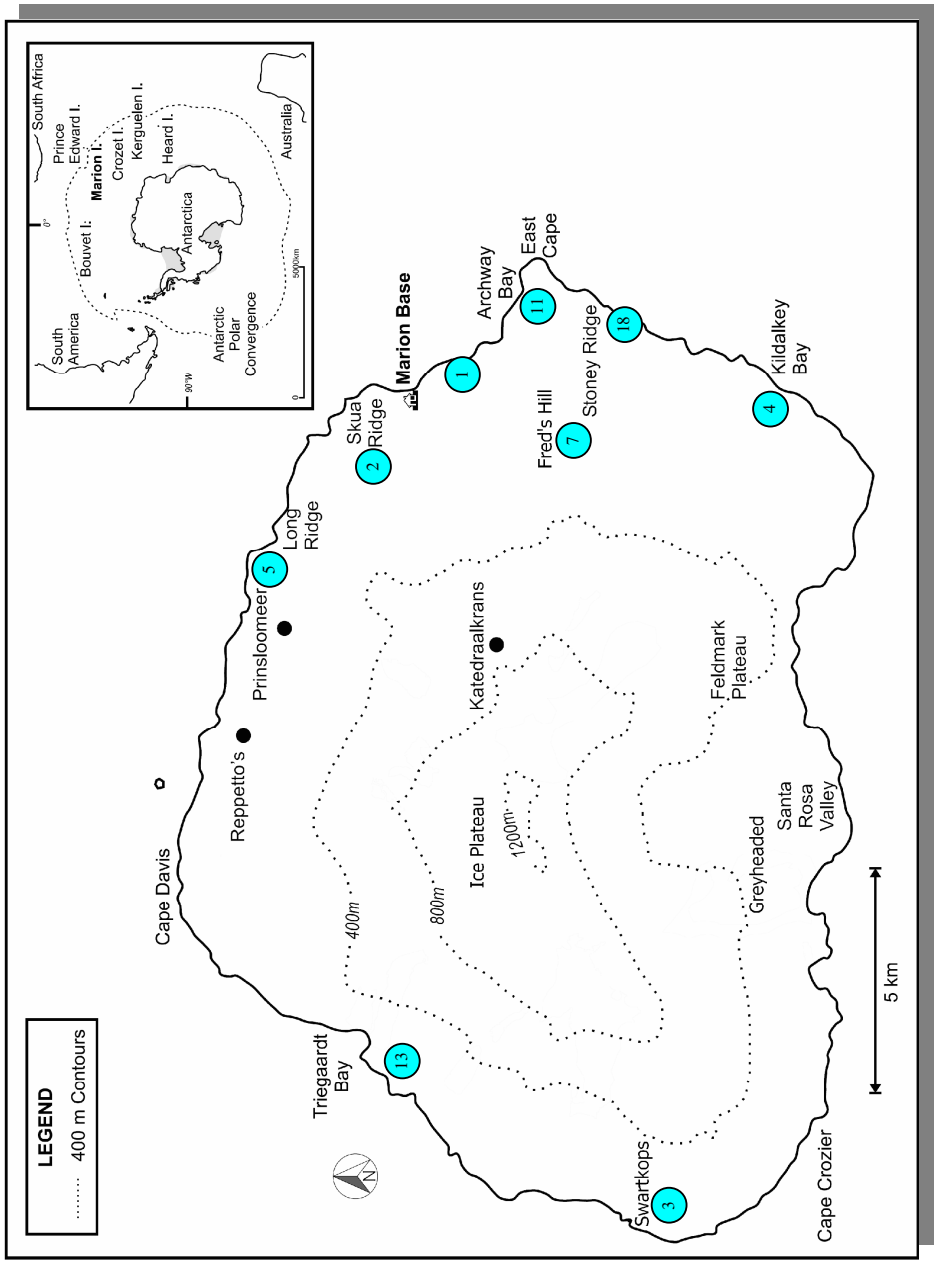

Figure 3.11: Map of Marion Island redrawn from Nel *et al*, 2003. Circles indicate the sampling localities for *I.* cf. *palustris.* Locality numbers correspond to those given in Table 2.1. There are no differences between any of the sampling localities.

3.3.2 COII gene

Five hundred and seventy base pairs were analysed for 42 individuals from the same localities as for COI (see Figure 3.11). The base frequencies calculated in PAUP* (Swofford, 2000) for A:C:G:T were 0.28: 0.25: 0.16 and 0.31 respectively. Two haplotypes were detected for the 42 *I.* cf *palustris* specimens included for the COII gene. A single specimen, sampled at Swartkops (3) differed by one mutational step from the remainder of the specimens sampled across the island. Haplotype diversity was calculated at 0.048 and nucleotide diversity (as calculated by DnaSP version 3, Rozas & Rozas, 1999) was 0.00016.

3.3.3 Combined data set

Thirty-nine specimens included in this part of the study had comparable data sets for both COI and COII. The base frequencies calculated in PAUP* (Swofford, 2000)for A:C:G:T were 0.28: 0.22: 0.20: 0.28, respectively. A single specimen bore a haplotype different from the shared common haplotype separated by one mutational step.

For both the COII data set and the combined data, there was only a single variable position and no parsimony-informative sites. Arlequin version 2 (Schneider *et* al, 2000) revealed an Fs-value of –7.54 ($p<0.001$) for the combined data set, and the mismatch distribution (Figure 3.12) shows that the observed population data fit exactly to the expected values of a dynamic population under the assumption of sudden expansion. As there was no divergence, neither the Mantel test nor nested cladistic analysis was done for this species. A minimum spanning network was constructed, using calculations by Arlequin for minimum number of mutational steps. This shows the one central haplotype representing all but one individual, which was sampled at Swartkops (3) (Figure 3.13).

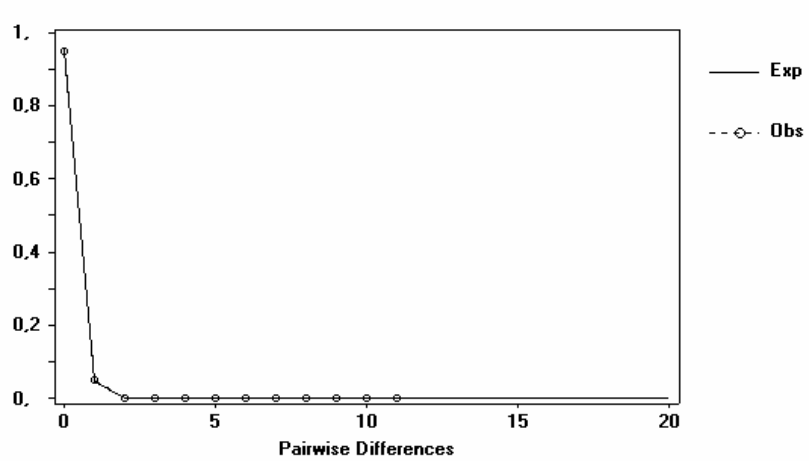

Figure 3.12: Mismatch distribution with the number of differences separating haplotypes plotted against their respective frequencies for 1050 bp of the combined (COI and COII) data set for 39 *I. palustris* specimens. The solid line indicates the expected values and the broken line the obesreved values.

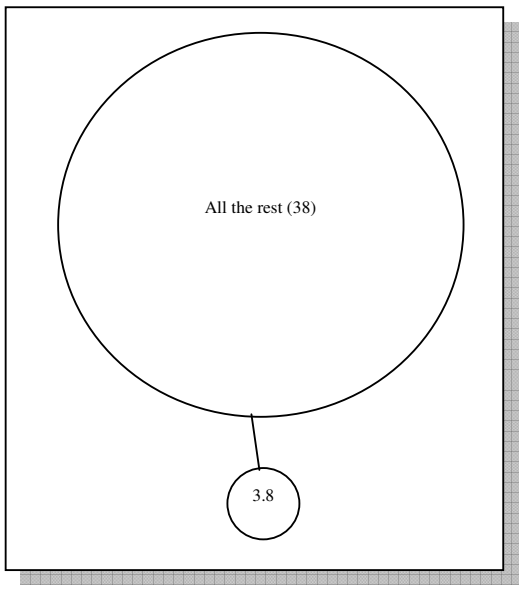

Figure 3.13: Minimum spanning network based on 1050 bp of the combined data set for 39 individuals of. *I. palustris*. Lines represent one mutation difference between haplotypes. Circles representing haplotypes are drawn according to scale with specimens characterised by haplotypes within (except where a large number of specimens shared a haplotype in which case only the number of specimens is indicated).

Chapter 4

Discussion

Discussion

Mitochondrial sequence variability for *C. antarcticus, T. bisetosa* and *I.* cf. *palustris*

The proportion A+T in the COI gene for the three species included in this study ranged between 55% and 61%. These values are the same as those reported for the same gene in other Collembola species (63%-66%; Frati *et al*, 2000; Soto-Adames, 2002; Stevens & Hogg, 2003), but much lower than that reported for other insects such as Lepidoptera (91% in Vandewoestijne *et al*, 2004). The A+T bias for the COII gene (45%-59%) for the three species included in the present study, is more in line with previously published results in Collembola (64%-66%; Frati & Dell'Ampio, 2000; Frati *et al*, 2001). When considering the COI and COII genes in concert, the A+T bias (56%-61%) is lower than that calculated for the entire mitochondrial genome of another species of Collembola, *Tetrodontophora bielensis* (73% in Nardi *et al*, 2001). This could be the result of the A+T rich region not being included in our sequences. From an evolutionary perspective it is interesting that the A+T bias in Collembola in general is noticeably lower than those reported for other insects such as *Drosophila yakuba* (79%), honey bees (85%) and butterflies (92%) (Crozier & Crozier, 1993; Vandewoestijne *et al*, 2004). The structure and heteroplasmy of the A+T rich region is used in phylogentic studies to determine genetic relationships among species and although speculative, this difference in A+T bias might be due to the fact that Collembola is more closely related to Crustacea (Nardi *et al*, 2003) or it may in fact reflect the primitive state for insects. The structure and heteroplasmy of the A/T rich region is used in phylogentic studies to determine genetic relationships among species.

The overall interspecific sequence divergence within *C. antarcticus* (COI: 0.18-1.32%; COII: 0.25-1.22%) and *T. bisetosa* (COI: 0.13%-0.83%) is similar to what was reported in other studies done on Collembola species in Antarctica for both COI (0.2%-2% in Stevens & Hogg, 2003) and COII (0.15%-1.9% in Frati *et al*, 2001). It is interesting to note that divergence values for species included in this study is much lower than for a saproxylic Collembola in New Zealand, (Garrick, 2002) where COI shows up to 6% sequence divergence. This observation could possibly be ascribed to *C. antarcticus* and *T. bisetosa* not being as habitat-specific as this saproxylic species (Gabriel *et al*, 2001; Garrick *et al*, 2004).

Phylogeography of indigenous species *C. antarcticus* and *T. bisetosa*

Marion Island is generally believed to be between 500 000 and 1 million years old (Smith, 1987; Hänel & Chown, 1998); however, the latest estimates based on K-Ar dating place closer to 450 000 years old (McDougal *et al*, 2001). Climatic and geomorphological evidence suggest that during the last 450 000 years Marion Island has been exposed to approximately five glacial and eight volcanic episodes. The three most recent volcanic episodes (VI, VII and

VIII, as named in McDougal *et al*, 2001) corresponded with interglacial stages while episodes I, II and IV coincided with major glaciations (McDougal *et al*, 2001). The remaining two volcanic episodes, III and V, are not as well documented and it is not clear whether they coincided with glacial or interglacial stages (McDougal *et al*, 2001).

Two types of lava are present on Marion Island; the older grey lava is dated between 450 000 ±80 000 and 23 000 ± 7 000 and shows evidence of glaciation. The younger black lava (dated to 17 000 ± 8000 years) does not show signs of glaciation, implying that these black lava flows would have occurred subsequent to the last glaciation episode (McDougal *et al*, 2001). From a population perspective these lava flows are important since the younger black lava over the older grey layers would have reinforced existing barriers and perhaps even created new barriers separating populations. Important also is that these areas of lava flow present uninhabitable regions in themselves, possibly for long periods of time.

A molecular clock was applied to estimate whether these climatic changes influenced the divergence in the two indigenous species included in the present study. Despite problems using molecular clocks to estimate divergence times between organisms (Alroy, 1999; Lee 1999; Conroy & Van Tuinen, 2003; Graur & Martin, 2004) there is no alternative in the absence of reliable fossil data to use as calibration points. Despite obvious problems, we followed Soto-Adames (2002) and employed a molecular rate of 2.3% per million years to our data. It should be noted that these estimates of divergence times are only rough estimates based on the percentage sequence divergence between populations and typically have large standard errors associated with them. Our only aim is to provide some idea of the relative ages of populations across the island.

For *C. antarcticus*, divergence time estimates obtained from the COI and COII genes were remarkably similar with the divergence times for the younger populations being less than 100 000 years and the oldest populations being separated for approximately 500 000 years (COI: 90 000-572 000 years (±115 500 years); COII: 96 000-528 400 years (±81 000 years)). Assuming that Marion Island is indeed only approximately 500 000 years old (McDougal *et al*, 2001), these older estimates would indicate that *C. antarcticus* colonised Marion Island shortly after its formation and that this species may in fact represent one of the first colonists to the island.

Based on COI gene estimates, *T. bisetosa* is a much later colonist to Marion Island, with population divergence times raging between 57 700 years for the more recently diverged and 360 000 years (±57 000 years) for the oldest populations.

For Collembola to have survived harsh conditions described above (i.e, repeated volcanic eruptions and glaciation cycles), certain regions must have remained relatively ice-free. These would have acted as refugia where Collembola survived and from where they recolonised available habitat once climatic conditions on the island became more favourable. Evidence for such refugia comes from flowering plants and ferns on Marion Island which were present 16 000 years ago (Van Zinderin Bakker *et al*, 1966) suggesting that the island, in its entirety, was not glaciated at that time. Also, Nel *et al*, 2003 indicated that the Feldmark Plateau was either ice-free or deglaciated earlier than the rest of Marion Island suggesting that there were indeed ice-free refugia present on the island.

Molecular data further support such refugia across Marion Island typically characterised by high intra-population divergences. For *C. antarcticus*, the highest intra-population divergences for both the COI and COII genes are reported from Katedraal (9) (COI: 0.80 ± 0.36%; COII: 1.22 ± 0.53%; see Tables 3.1 and 3.5), a locality situated at high altitude close to the ice-plateau and permanent glaciers. Furthermore, some of the most divergent haplotypes are found at this locality indicating the relatively old age of this population (as opposed to a recent colonisation). It is speculated that this elevated region would have experienced many glaciation events (perhaps more so than mid- and low-altitude localities where the ocean might have prevented permanent glaciations) yet that certain areas (for example Katedraal) remained (relatively) ice-free where fauna and flora survived.

With shifts from glacial to interglacial stadia, the worldwide trend in terrestrial ecosystems was to expand their ranges (Graham *et al*, 2003). Fu's Fs for COI in *T .bisetosa* (-8.728; $p<0.001$) and for the same gene in *C. antarcticus* (-26.04; $p<0.001$) are indicative of populations that are / were in disequilibrium. Mismatch distribution for both indigenous species show a very good fit to the expected values for a population under the model of a sudden expansion and there is a strong signature of population expansion evident from the haplotype networks. When taken together, population expansions are put forward as a possible explanation to account for the observed disequilibrium. It is important to note that we did not test for selection or migration and that both of these could also cause populations to deviate from equilibrium. However, the results from the mismatch distribution would suggest that Collembola populations have expanded on Marion Island. This makes ecological and evolutionary sense given the history of glacial and interglacial periods.

When considering the haplotype networks for *C. antarcticus* (both for the COI gene based on 75 specimens (Figure 3.3) and for the combined COI and COII genes for 44 specimens (Figure 3.7), two distinct population signatures are evident. First, it shows a pattern that is typical of a recent population expansion followed by a stasis period allowing for few

differences to accumulate. Secondly, there is also a deeper and more ancient pattern where haplotypes are not shared by as many specimens and occur more as singletons with higher mutational differences separating specimens. This would be typical where divergence times have been sufficiently long allowing for differences to accumulate between specimens (Templeton, 1995; 1998). When the two genes are considered in concert, the results suggest an old population / species with more recent demographic expansions superimposed on top of it. This is corroborated by significant Fu's Fs values (-26.04; $p<0.001$ for COI and -26.61; $p<0.001$ for COII).

When considering the haplotype network for *T. bisetosa* based on 40 specimens for the COI gene, the emerging pattern is somewhat different. Although there are still central haplotypes shared by many specimens, the pattern of recent expansion is less distinct. Three (central) haplotypes are shared by 15, 7 and 7 specimens with few singletons connecting directly to it. The pattern for *T. bisetosa* is one of higher population stability rather than very recent expansion as is the case for *C. antarcticus*. Notwithstanding that fewer specimens as well as shorter sequences are considered for *T. bisetosa*, it is interesting to correlate these patterns to the species' respective life history strategies. *C. antarcticus* is a hemiedaphic species and therefore has an elevated degree of uncertainty in life-strategy, as the environment they live in is less constant and sheltered than that of the soil-living species (Peterson, 1980, 2002). In comparison, *T. bisetosa,* which is a euedaphic species living under conditions that are more buffered to climate changes. Taking into account only this aspect of the life histories of these two species, one would expect *C. antarcticus* to experience a more unstable evolutionary history characterized by population bottlenecks and confinement to refugia. In contrast, *T. bisetosa* would be faced with fewer evolutionary uncertainties and thus be characterized by a more stable population structure. The fact that *T. bisetosa* seems to have more stable living conditions could also contribute to the population being closer to equilibrium than *C. antarcticus*, as the unpredictable conditions at the surface would require a high degree of genetic variation (Peterson, 2002).

Contrary to studies on *Orchesella cincta* that occur in forests in the Netherlands (Van der Wurff *et al*, 2003) and *Gressittacantha terranova* in Antarctica (Fanciulli *et al*, 2001), no correlation was observed between genetic and geographical distances. The lower-lying parts of Marion Island have a more favourable climate when compared to the higher lying parts which are around the permanent ice cap. It is therefore expected that a collembolan would travel around the island rather than across the central ice cap. As such, geographical distances used to test for isolation by distance were calculated along these proposed routes of colonisation. Although the lack of positive correlation may be due to misjudgement concerning these proposed routes of colonisation, we do not believe that this is the case. Little

or no geographic structuring of genetic variation can also be invoked to explain the lack of isolation by distance (Vandewoestijne *et al*, 2004). For both *C. antarcticus* and *T. bisetosa* overall population F_{ST}'s were significant albeit low (see Tables 3.2, 3.6, 3.8 and 3.11). Furthermore, a nested clade analysis for *C. antarcticus* indicated that certain clades contained haplotypes that are only found in particular geographical regions for both COI and the combined data set. Therefore, we propose that the lack of isolation by distance in these species is a true reflection of genetic structure within the species although this partitioning of variation is weak and shallow (i.e. not deep genetic differences separating populations / regions). Although *C. antarcticus* is considered one of the ancient colonisers of the island, it is not expected that they would show very deep genetic divergences as the island itself is relatively young.

Kildalkey Bay (4) is a noteworthy locality. Despite there being no obvious (human-perceived) evidence of barriers to gene flow between Kildalkey Bay and the rest of the island, there seems to be a distinctly higher level of genetic divergence between this locality and the rest of the island. Without fail, SAMOVA analyses for both *C. antarcticus* and *T. bisetosa* singled out Kildalkey Bay as distinct from the rest (irrespective of the number of groups defined). In addition, when considering pairwise F_{ST} values between populations, most of the significant values were between Kildalkey Bay and remaining localities across the island.

It appears that most of the new arrivals (including a moth and woody plant; pers. comm. Sarette Slabber) onto Marion Island make a first appearance at Kildalkey Bay (4) rather than at the scientific base which would seem the most likely port of entry due to the relatively extensive human activity at this point. This has resulted in speculation that although the island is dominated by a predominantly westerly wind (Smith, 1987), animals and seeds that arrive naturally on the island mostly make their first arrival at Kildalkey Bay and spread across the island from here (personal communication, Steven Chown, Richard Mercer).

This argument is strengthened by the fact that diaspora and marine plankton from Marion and Prince Edward Islands tend to show a higher taxonomic affinity to the Crozet and Kergeulen archipelagos, which would mean that migration took place against the direction of the prevailing westerly winds and ocean currents (Van Zinderen Bakker *et al*, 1966). A similar situation is found on another sub-Antarctic island, Macquarie Island, where travelling depressions on the island give rise to winds other than from the west and mite species are then dispersed from east to west, against the prevailing west-to-east air and water currents (Watson, 1967; Wallwork, 1973). It should be noted though that the glacial-interglacial cycles of the quaternary have driven long-term changes in oceanographic conditions (Graham *et al*, 2003) and the direction of the currents may also have changed during these times explaining

why Kildalkey Bay (4) might have been the first point of contact for organisms arriving on Marion Island.

Comparison between indigenous and introduced species

Long term studies in the U.S desert shrubland and pinyon-juniper woodlands revealed effects that weren't detected in shorter experiments and revealed that changes in climate, and other abiotic factors, can have nonlinear consequences as they are augmented by species interactions such as competition (Brown *et al*, 2001). Extreme environmental circumstances can cause apparently inconsequential organisms to have unexpectedly large effects on an ecosystem. In one study (Brown *et al*, 2001), a moth that in one set of conditions is merely one more herbivore causing little damage to pinyon, has in another set of conditions come forward as a keystone species with large, cascading effects on the ecosystem.

Sub-Antarctic islands are rather isolated and show a depauperate invertebrate fauna in comparison to continental habitats, resulting in lower levels of colonisation and replacement within such ecosystems. This, combined with harsh environmental conditions, makes these islands very sensitive to disturbances (Brown *et al*, 2001; Smith, 2002). Species on islands evolve in seclusion as opposed to those on continents where competition is often fierce for resources. These island species can be more vulnerable to aggressive invaders. Alien species often have a greater effect on island species, especially on Darwinian islands where neo-endemics occur (Gillespie & Roderick, 2002).

On another sub-Antarctic island, South Georgia, it has been noted that an introduced species, *Hypogastrura purpurescens*, has already displaced *C. antarcticus* from certain ranges of its habitat. This can be seen in the fact that *C. antarcticus* occurs in certain types of habitat in specific places on the island where there has been limited human contact. In places where this contact is more prominent, *H. purpurescens* occurs, instead of *C. antarcticus* (Convey *et al*, 1999).

Gabriel *et al* (2001) found that there was no evidence of negative interactions between introduced and indigenous species, but given time and the rapid temperature changes the Southern Ocean is currently experiencing (Bergstrom & Chown, 1999; Smith, 2002), this is certain to change. The difference in distribution patterns of indigenous and exotic Collembola on Marion Island indicates that the occurrence of introduced insect species on sub-Antarctic islands is related to mean annual temperature and human occupation (Chown *et al*, 2002). Introduced species have been shown to occur more frequently in the warmer, dryer habitats (Gabriel *et al*, 2001). Over the last 50 years, the mean annual temperature on Marion Island

has risen by more than 1˚C and annual precipitation has declined by ~600 mm (Smith, 2002). If this trend continues, introduced species may displace indigenous ones.

Although Gabriel *et al* (2001) found that introduced species of Collembola don't seem to have a negative effect on the indigenous Collembolan fauna, the difference in the distribution of the two groups may very well be considered evidence that negative competition is indeed the case, forcing endemic taxa to move away from the warmer moist lowland habitats (Chown *et al*, 2002). *Isotomurus.* cf *palustris* has been able to spread into these habitats without the necessity to adapt genetically as measured by nearly neutral mtDNA markers (with the exception of a single specimen, only a single mitochondrial haplotype characterize 44 specimens for the COII gene and 39 specimens for the combined data set included in our study), and as the temperature in the Southern Ocean rises and the air temperatures warm up, these aliens are expected to have the capacity to spread into the warming high-altitude habitats on the island as well and may eventually out compete the indigenous species.

On the whole, the introduced invertebrates on Marion seem to grow and reproduce faster than indigenous ones, completing more than one life-cycle in a year. By comparison, the indigenous species have extended life-cycles of a year or more (Barendse, 2000; Chown *et al*, 2002). Although temperature changes will affect both introduced and indigenous species, the introduced species, with their faster generation times, may be at an advantage and may be able to adapt faster (Barendse, 2000; Chown *et al*, 2002).

Species range extensions require selection as well as adaptation to different environments and the biota that is already present in the area (Hewitt, 2000); this adaptation normally takes an extensive amount of time. If conditions are already favourable for certain invasive organisms, they would have an "unfair" advantage over indigenous taxa. The indigenous biota of Marion Island, which have evolved under cool moist conditions and are adapted to them, will be negatively affected if the island continues to show this warming trend. This will also facilitate the ease with which future introduced species can colonise the island (Bergstrom & Chown, 1999; Smith, 2002) and possibly become invasive. Warming of the Southern Ocean is concentrated in the sub-Antarctic Front and is occurring at a rate nearly double the global trend (Gille, 2002).

Isotomurus cf. *palustris* arrived on Marion Island a mere 25 years ago (Déharveng, 1981), and has since then managed to spread across the island, with the exception of only the highest and coldest habitats (Gabriel *et al*, 2001). Indigenous and introduced Collembola species still occupy different ecological niches at this point in time (Gabriel *et al*, 2001) and as of yet, *I.* cf. *palustris* has not diverged genetically from the haplotype that came onto the island 25 years ago. This species has thus far not extended it's range to the colder parts of Marion

Island. As the Southern Ocean warms up and climate change makes other habitats more accessible to this species, it may well displace the indigenous species from their ranges and become a very real threat to the indigenous Collembola fauna on Marion Island as well as the other sub-Antarctic islands to which *I* cf. *palustris* has managed to spread.

Conservation suggestions for Marion Island

The findings from our work have implications beyond simply describing the population structure for three species of Collembola:

First, our work has indicated the presence of refugia across Marion Island, most noticeably the locality of Katedraal, and specifically for *C. antarcticus*. This locality, which is situated close to the ice plateau, is speculated to have remained relatively ice-free during periods of extensive glaciation on Marion Island (Nel *et al*, 2003) allowing remnant populations to survive these harsh environmental conditions. It is of critical importance to have a conservation plan in place that takes areas with high genetic variance and possible past refugia into account. Our study was limited to only two indigenous species and further studies on indigenous species are needed to elucidate the exact position of these refugia. However, we strongly urge that a precautionary approach be taken whereby these areas of high genetic variability should receive special conservation attention.

Secondly, Kildalkey Bay was repeatedly shown to have larger divergences between localities when comparisons are made between all localities, suggesting that this is the landing point of organisms that arrive naturally on Marion Island. Above-ground and below-ground components of an ecosystem are linked and have different effects on each other, depending on the context in which they are interacting. (Wardle *et al*, 2004). "Little is understood about how invasion of soil organisms influences aboveground biota, although these effects should be strongest when the invading species has functional attributes that are not shared by the resident indigenous species" (Wardle *et al*, 2004). In a study on ascidians it was found that with a temperature difference of merely 3ºC, introduced species increased at double the rate of native species (Stachowicz *et al*, 2002). Invading soil invertebrates such as Collembola may therefore have a negative effect on the plant communities in the long run.

The third point concerns introduced species on Marion Island. Although the influence of introduced and invasive species are well documented for continents (Convey *et al*, 1999; Brown *et al*, 2001; Stachowicz *et al*, 2002) little is known about the effects of these intruders on sensitive sub-Antarctic island ecosystems. We have already established that at *I.* cf. *palustris* has the ability to spread rapidly (Gabriel *et al*, 2001) and with further changes in climate, it is almost certain (based on evidence from elsewhere (Convey *et al*, 1999) where

indigenous species were displaced) to outcompete endemic species for resources. It is imperative that further introductions are prevented.

References

Alroy J (1999) The fossil record of North American mammals: evidence for a Paleocene evolutionary radiation. *Systematic Biology*, **48**, 107-118.

Alpers DL, van Vuuren BJ, Arctender P, Robinson, T (2004) Population genetics of the roan antelope (*Hippotragus equinus*) with suggestions for conservation. *Molecular Ecology*, **13**, 1771-1784.

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics,* **18**, 489-522.

Ballard JWO, Whitlock M (2004) The incomplete natural history of mitochondria. *Molecular Ecology*, **13**, 729-744.

Barendse J (2000) Patterns in the distribution and abundance of terrestrial arthropods on sub-Antarctic Marion Island. *MSc Thesis.* University of Pretoria.

Barendse J, Chown SL (2001) Abundance and seasonality of mid-altitude fellfield arthropods from Marion Island. *Polar Biology,* **24**, 73-82.

Bergström DM, Chown SL (1999) Life at the front: history, ecology and change on Southern Ocean islands. *Trends in Ecology and Evolution*, **14**, 472-477.

Berry O, Gleeson DM (2004) Distinguishing historical fragmentation from a recent population decline - shrinking or pre-shrunk skink from New Zealand? *Biological Conservation,* **123**, 197-210.

Block W (1985) Arthropod interactions in an Antarctic terrestrial community. *British Antarctic Survey Bulletin,* 614-619.

Blouin MS, Yowell CA, Courtney CH, Dame JB (1995) Host movement and the genetic structure of populations of parasitic nematodes. *Genetics*, **141**, 1007-1014.

Boelhouwers J, Holness S, Sumner P (2002) The maritime Subantarctic: a distinct periglacial environment. *Geomorphology*, **1260**, 1-17.

Broady PA (1979) Feeding studies on the collembolan *Cryptopygus antarcticus* Willem at Signy Island, South Orkney Islands. *British Antarctic Survey*, Bulletin **48**, 37-46.

Brown JH, Whitham TG, Ernest SKM, Gehring CA (2001) Complex species interactions and the dynamics of ecological systems: long-term experiments. *Science*, **293**, 643-650.

Brown JM, Abrahamson WG, Way PA (1996) Mitochondrial DNA phylogeography of host races of the goldenrod ball gallmaker, *Eurosta solidaginis* (Diptera: Tephritidae). *Evolution*, **50**, 777-786.

Campbell MK (1999) *Biochemistry*. Harcourt Brace & Company, USA, 550-551.

Carapelli A, Frati F, Fanciulli PP, Dallai R (1995) Genetic differentiation of six sympatric species of *Isotomurus* (Collembola, Isotomidae); is there any difference in their microhabitat preference? *European Journal of Soil Biology*, **31**, 87-99.

Carapelli A, Frati F, Fanciulli PP, Dallai R (1997) Estimating levels of gene-flow between natural populations in the genus *Isotomurus* (Collembola, Isotomidae). *Pedobiologia*, **41**, 19-25.

Carapelli A, Frati F, Fanciulli PP, Nardi F, Dallai R (2004) Assessing species boundaries and evolutionary relationships in a group of south-western European species of *Isotomurus* (Collembola, Isotomidae) using allozyme data. *Zoologica Scripta*, **34**, 71-79.

Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. *Evolution*, **21**, 550-570.

Chown SL, McGeoch MA, Marshall DJ (2002) Diversity and conservation of invertebrates on the sub-Antarctic Prince Edward Islands. *African Entomology*, **10**, 67-82.

Conroy, CJ, Van Tuine M (2003) Extracting time from phylogenies: positive interplay between fossil and genetic data. *Journal of Mammalogy*, **84**, 444-455.

Convey P, Greenslade P, Arnold RJ, Block W (1999) Collembola of the sub-Antarctic South Georgia. *Polar Biology*, **22** 1-6.

Coulson SJ, Hodkinson ID, Webb NR, Harrison JA (2002) Survival of terrestrial soil-dwelling arthropods on and in seawater: implications for trans-oceanic dispersal. *Functional Ecology*, **16**, 353-356.

Creer S, Malhorta A, Thorpe RS, Chou WH (2001) Multiple causation of phylogeographical pattern as revealed by nested clade analysis of the bamboo viper (*Trimeresurus stejnegeri*) within Taiwan. *Molecular Ecology*, **10**, 1967-1981.

Crozier RH, Crozier YC (1993) The complete mitochondrial genome of the honeybee *Apis mellifera*: complete sequence and genome organization. *Genetics*, **133**, 97-117.

Cruickshank RH (2002) Molecular markers for the phylogenies of mites and ticks. *Systematic and Applied Acarology*, **7**, 3-10.

Déharveng L (1981) Collemboles des Iles subantarctiques de l'Ocean Indien Mission J. *Trave 1972-1973- Comite National Français Recherches Antarctiques*, **48**, 33-107.

Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology*, **11**, 2571-2581.

Edmonds CA, Lillie AS, Cavalli-Sforza LL (2004) Mutations arising in the wave front of an expanding population. *Proceedings of the National Academy of the Sciences*, **101**, 975-979.

Emerson BC (2002) Evolution on oceanic islands: molecular phylogenetic approaches to understanding pattern and process. *Molecular Ecology*, **11**, 951-966.

Excoffier L, Smouse PE, Quattro JM (1992) Analyses of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479-491.

Fanciulli PP, Melegari D, Carapelli A, Frati F, Dallai R (2000) Population structure, gene flow and evolutionary relationships in four species of the genera *Tomocerus* and *Pogonognathellus* (Collembola, Tomoceridae). *Biological Journal of the Linnean Society*, **70**, 221-238.

Fanciulli PP, Summa D, Dallai R, Frati F (2001) High levels of genetic variability and population differentiation in *Gressittacantha terranova* (Collembola, Hexapoda) from Victoria Land, Antarctica. *Antarctic Science*, **13**, 246-254.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for the amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, **3**, 294-299.

Frati F, Dell'Ampio E (2000) Molecular phylogeny of three subfamilies of the Neanuridae (Insecta, Collembola) and the position of the Antarctic species *Friesae grisea. Pedobiologia*, **44**, 343-355.

Frati F, Dell'Ampio E, Carapelli A, Fanciulli PP, Nardi F, Spinsati G, Dallai R (2000) DNA sequence analysis to study the evolution of Antarctic Collembola. *Italian Journal of Zoology SUPPLEMENT*, **1**, 133-139.

Frati F, Simon C, Sullivan J, Swofford DL (1997) Evolution of the mitochondrial cytochrome oxidase II gene in Collembola. *Journal of Molecular Evolution*, **44**, 145-158.

Frati F, Spinsati G, Dallai R (2001) Genetic variation of mtCOII gene sequences in the Collembolan *Isotoma klovstadi* from Victoria Land, Antarctica: evidence for population differentiation. *Polar Biology,* **24**, 934- 940.

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915-925.

Fuselli S, Tarazona-Santos E, Dupanloup I, Soto A, Luiselli D, Pettener D (2003) Mitochondrial DNA diversity in South America and the genetic history of Andean Highlanders. *Molecular Biology and Evolution*, **20**, 1682-1691.

Gabriel AGA, Chown SL, Barendse J, Marshall DJ, Mercer RD, Pugh PJA, Smith VR (2001) Biological invasions of Southern Ocean islands: the Collembola of Marion Island as a test of generalities. *Ecography*, **24**, 421-430.

Gaggiotti OE, Brooks SP, Amos W, Harwoods J (2004) Combining demographic, environmental and genetic data to test hypotheses about colonisation events in metapopulations. *Molecular Ecology*, **13**, 811-825.

Garrick R (2002) Phylogeography of a new species of Collembola (Neanuridea) from Tallaganda State Forest, New South Wales. *M.Sc. Thesis,* Australian National University.

Garrick R, Sands CJ, Rowell DM, Tait NN, Greenslade P, Sunnucks P (2004) Phylogeography recapitulates topography: very fine-scale local endemism of a saproxylic "giant" springtail at Tallaganda in the Great Dividing Range of south-east Australia. *Molecular Ecology*, **13**, 3329-3344.

Gille ST (2002) Warming of the Southern Ocean since the 1950's. *Science*, **295**, 1275-1277.

Gillespie RG, Roderick GK (2002) Arthropods on islands: colonisation, speciation and conservation. *Annual Revue of Entomology*, **47**, 595-632.

Gisin H (1943) Őkologie und Lebensgemeinschaften der Collembolan im schweizerischen Exkursionsgebiet Basels. *Revue Suisse de Zoologie*, **50**, 131-224.

Graham MH, Dayton PK, Erlandson JM (2003) Ice ages and ecological transitions on temperate coasts. *Trends in Ecology and Evolution*, **18**, 33-40.

Graur D, Martin W (2004) Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision. *Trends in Genetics*, **20**, 80-86.

Gremmen NJM (1981) The vegetation of the Subantarctic islands Marion and Prince Edward. Dr. W. Junk, The Hague.

Hale LR, Singh RS (1991) A comprehensive study of genetic variation in natural populations of *Drosophila melanogaster*. IV. Mitochondrial DNA variation and the role of history *vs.* selection in the genetic structure of geographic populations. *Genetics*, **129**, 103-117.

Hall KJ (1990) Quaternary glaciations in the Southern Ocean: sector 0º long -180º long. *Quaternary Science Reviews,* **9**, 217-228.

Hall KJ (2002) Review of present and Quaternary peri-glacial processes and landforms of the maritime and sub-Antarctic region. *South African Journal of Science,* **98**, 71-81.

Hänel C, Chown S (1998) An introductory guide to the Marion and Prince Edward Island special nature reserves. *Department of Environmental Affairs and Tourism, First Edition*.

Hanski IA, Gilpin ME (1997) Metapopulation biology: ecology, genetics and evolution. *Academic Press, New York*.

Harpending HC, Batzer MA, Gurven M, Jorde LB, Rodgers AR, Sherry ST (1998) Genetic traces of ancient demography. *Proceedings of the National Academy of Science,* **95**, 1961-1967.

Hedges SB, Kumar S (2004) Precision of molecular time estimates. *Trends in Genetics*, **20**, 242-247.

Hedrick PH (1999) Perspective: highly variable loci and their interpretation in evolution and conservation. *Evolution*, **53**, 313-318.

Hewitt GM (1996) Some genetic consequences of the ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58**, 247-276.

Hewitt GM (2000) The genetic legacy of the Quarternary ice ages. *Nature*, **405**, 907-908.

Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*, **167**, 747-760.

Hillis DM, Moritz C, Mable BK (1996) *Molecular Systematics*, 2nd edition. Sinauer Associates, Sunderland, MA.

Hopkin SP (1997) *Biology of the springtails (Insecta: Collembola)*. Oxford University Press. Oxford.

Kimura M, Weiss G (1964) The stepping-stone model of population structure and the decrease of genetic correlation with distance. *Genetics*, **49**, 561-576.

King TM, Kennedy M, Wallis GP (2003) Phylogeographic genetic analysis of the alpine weta, *Hemideina maori*: evolution of a colour polymorphism and origins of a hybrid zone. *Journal of the Royal Society of New Zealand*, **33**, 715-729.

Kingman JFC (1980) Mathematics of genetic diversity. *Society for Industrial and Applied Mathematics*, Philadelpia.

Knowles LL (2000) Tests of Pleistocene speciation in montane grasshoppers (genus *Melanoplus*) from the sky islands of Western North America. *Evolution*, **54**, 1337-1348.

Knowles LL, Maddison WP (2002) Statistical phylogeography. *Molecular Ecology*, **11**, 2623-2635.

Lee MSY (1999) Molecular clock calibrations and metazoan divergence dates. *Journal of Molecular Evolution*, **49**, 385-391.

Lesia MGA, Hofmeyr MD, D'Amato ME (2003) Genetic variation in three *Chersina angulata* (angulate tortoise) populations along the west coast of South Africa. *African Zoology,* **38**, 109-117.

Li M, Wei F, Groves P, Feng Z, Hu J (2003) Genetic structure and phylogeography of the takin (*Budorcas taxicolor*) as inferred from mitochondrial sequences. *Canadian Journal of Zoology*, **81**, 462-468.

Liu H, Beckenbach AT (1992) Evolution of the mitochondrial cytochrome oxidase II gene among 10 orders of insects. *Molecular Phylogenetics and Evolution*, **1**, 41-52.

Lloyd B (2003) The demographic history of the New Zealand short-tailed bat *Mystacina tuberculata* inferred from modified control region sequences. *Molecular Ecology*, **12**, 1895-1911.

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning. Cold Spring Harbor, New York.

McDougal I, Verwoerd W, Chevallier L (2001) K-Ar geochronology of Marion Island, Southern Ocean. *Geological Magazine*, **138**, 1-17.

Mekawa K, Kitade O, Matsumoto T (1999) Molecular phylogeny of orthopterid insects based on the mitochondrial cytochrome oxidase II gene. *Zoological Science*, **16**, 175-184.

Miller SA, Harley JP (1996) *Zoology*, 3rd edition. Wm. C. Brown Publishers, London.

Miura T, Roisin Y, Matsumoto T (2000) Molecular phylogeny and biogeography of the nasute termite *Gebus nasutitermites* (Isoptera: Termitidae) in the Pacific tropics. *Molecular Phylogenetics and Evolution*, **17**, 1-10.

Nardi F, Carapelli A, Fanciulli PP, Dallai R, Frati F (2001) The complete mitochondrial DNA sequence of the basal hexapod *Tetrodontophora bielensis:* evidence for heteroplasmy and tRNA translocations. *Molecular Biology and Evolution*, **18**, 1293-1304.

Nardi F, Spinsati G, Boore JL, Carapelli A, Dallai R, Frati F (2003) Hexapod origins: monophyletic or paraphyletic. *Science*, **299**, 1887-1889.

Near TJ, Meylan PA, Shaffer HB (2005) Assessing concordance of fossil calibration points in molecular clock studies: an example using turtles. *The American Naturalist*, **165**, 137-146.

Nel W, Holness S, Meiklejohn KI (2003) Observations on rapid mass movement and screes on sub-Antarctic Marion Island. *South African Journal of Science,* **99**, 177-181.

Nielson M, Lohman K, Sullivan J (2001) Phylogeography of the tailed frog (*Ascaphius truei*) : implications for the biogeography of the Pacific northwest. *Evolution*, **55**, 147-160.

Parker PG, Snow A.A, Schug MD, Booton GC, Fuerst PA (1998) What molecules can tell us about populations: choosing and using a molecular marker. *Molecular Techniques in Ecology*, **79**, (361-382).

Peterson H (1980) Population dynamic and metabolic characterization of Collembola species in a beech forest ecosystem. In: Dindal DL (ed) Soil biology as related to land use practices.

Proceedings of the VII International Soil Zoology Colloquium of the I.S.S.S., Syracuse, N.Y., U.S.A. July 29 Aug. 3, 1979. EPA 560/1380038, Washington D.C., 806-833.

Peterson H (2002) General aspects of collembolan ecology at the turn of the millennium. *Pedobiologia*, **46**, 246-260.

Pogson GH, Taggart CT, Mesa KA, Boutilier RG (2001) Isolation by distance in the Atlantic cod, *Gadus morhua* at large and small geographic scales. *Evolution*, **55**, 131-146.

Posada D (1998) PARSPROB v 1.1 Available at http://bioag.byu.edu/zoology/crandall_lab/programs.htm.

Posada D, Crandall KA, Templeton AR (2000) GEODIS: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Molecular Ecology*, **9**, 487-488.

Pugh PJA (1993) A synonymic catalogue of the Acari from Antarctica, the sub-Antarctic Islands and the Southern Ocean. *Journal of Natural History*, **27**, 323-421.

Ramos-Onsins SE, Rozas J (2002) Statistical properties of new neutrality tests against population growth. *Molecular Biology and Cell Genetics*, **44**, 198-208.

Rapoport EH (1971) The geographical distribution of neotropical and Antarctic Collembola. *Pacific Insects Monographs*, **25**, 99-118.

Ridley M (1993) *Evolution* (Chapter 16). Blackwell Science Inc, Oxford.

Roehrdanz RL, Degrugillier ME, Black WC (2002) Novel rearrangements of arthropod mitochondrial DNA detected with long-PCR: applications to arthropod phylogeny and evolution. *Molecular Biology and Evolution,* **19**, 841-849.

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution,* **9**, 552-569.

Rozas J, Rozas R (1999) DnaSP: an integrated program for molecular population genetics and molecular evolutionary analysis. *Bioinformatics,* **15**, 174-175.

Ruedi M, Castella V (2003) Genetic consequences of the ice ages on nurseries of the bat *Myotis myotis*: a mitochondrial and nuclear survey. *Molecular Ecology*, **12**, 1527-1540.

Schneider S, Excoffier L (1999) Estimation of past demographic parameters from the distribution of pairwise differences when mutation rates vary among sites: application to human mitochondrial DNA. *Genetics*, **152**, 1079-1089.

Schneider S, Roessli D, Excoffier L (2000) Arlequin: a software for population genetics data analysis. Ver2.000. Genetics and Biometry Laboratory, Department of Anthropology, University of Geneva.

Seddon JM, Santucci F, Reeve NJ, Hewitt G (2001) DNA footprints of European hedgehogs, *Erinaceus europaeus* and *E. concolor*: pleistocene refugia, postglacial expansion and colonisation routes. *Molecular Ecology*, **10**, 2178-2198.

Simmons RB, Weller SJ (2001) Utility and evolution of cytochrome b in insects. *Molecular Phylogenetics and Evolution*, **20**, 196-210.

Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P (1994) Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America*, **87**, 651-701.

Slatkin M (1985) Gene flow in natural populations. *Annual Review of Ecology and Systematics,* **16**, 393-430.

Slatkin M (1987) Gene flow and the geographic structure of natural populations. *Science*, **236**, 787-792.

Smith VR (1987) The environment and biota of Marion Island. *Suid-Afrikaanse Tydskrif vir Wetenskap*, **83**, 211-220.

Smith VR (2002) Climate change in the sub-Antarctic: an illustration from Marion Island. *Climatic Change*, **52**, 345-357.

Soto-Adames FN, (2002) Molecular phylogeny of the Puerto Rican *Lepidocyrtus* and *Pseudosinella* (Hexapoda: Collembola), a validation of Yoshii's "color pattern species". *Molecular Phylogenetics and Evolution*, **25**, 27-42.

Stachowicz JJ, Terwin JR, Whitlatch RB, Osman RW (2002) Linking climate change and biological invasions: ocean warming facilitates nonindigenous species invasions. *Proceedings of the National Academy of Science,* **99**, 15497-15500.

Stevens MI, Hogg ID (2003) Long-term isolation and recent range expansion from glacial refugia revealed for the endemic springtail *Gomphiocephalus hodgsoni* from Victoria Land, Antarctica. *Molecular Ecology*, **12**, 2357-2369.

Swofford DL (2000) PAUP*. Phylogenetic analyses using parsimony (*and other methods). Version 4.0b2a. Sinauer Associates, Inc., Publishers, Sunderland, Massachusetts.

Tajima F (1983) Evolutionary relationship of DNA sequences in finite populations. *Genetics*, **105**, 437-460.

Tamarin RH (1999) Principles of Genetics, 6th edition. Wm. C. Brown Publishers.

Templeton AR (1998) Nested clade analysis of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology,* **7**, 381-397.

Templeton AR (2004) Statistical phylogeography: methods of evaluating and minimizing inference errors. *Molecular Ecology*, **13**, 789-809.

Templeton AR, Boerwinkle E, Sing CF (1987) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. I. Basic theory and an analysis of alcohol dehydrogenase activity in *Drosophila*. *Genetics*, **117**, 343-351.

Templeton AR, Routman E, Phillips (1995) Separating population structure from population history: a cladistic analysis of the geographical dispersion of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. *Genetics*, **140**, 767-782.

Templeton AR, Sing CF (1993) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. IV. Nested analyses with cladogram uncertainty and recombination. *Genetics*, **134**, 659-669.

Tolley K, Groenewald JC, Gopal K, Matthee CA (2005) Mitochondrial DNA panmixia in spiny lobster *Palinurus gilchristi* supports a recent population expansion. *In press*

Travis JMJ, Dytham C (2002) Dispersal evolution during invasions. *Evolutionary Ecology Research*, **4**, 1119-1129.

Turner TF, Trexler JC, Harris JL, Haynes JL (2000) Nested cladistic analysis indicates population fragmentation shapes genetic diversity in a freshwater mussel. *Genetics*, **154**, 777-778.

Tzeng T, Yeh S, Hui C (2004) Population genetic structure of the kuruma prawn (*Penaeus japonicus*) in East Asia inferred from mitochondrial DNA sequences. *Journal of Marine Science*, **61**, 913-920.

Van der Wurff AWG, Isaaks A, Ernsting G, Van Straalen M (2003) Population substructures in the soil invertebrate *Orchesella cincta*, as revealed by microsattellite and TE-AFLP markers. *Molecular Ecology*, **12**, 1349-1359.

Van Zinderen Bakker EM, Winterbottom JM, Dyer RA (1966) Marion and Prince Edward Islands: report on the South African biological & geological expedition / 1965-1966. A.A. Balkema. Cape Town.

Vandergast AG, Gillespie RG, Roderick, GK (2004) Influence of volcanic activity on the population of Hawaiian *Tetragnatha* spiders: fragmentation, rapid population growth and the potential for accelerated evolution. *Molecular Evolution*, **13**, 1729-1743.

Vandewoestijne S, Baguette M, Brakefield PM, Saccheri IJ (2004) Phylogeography of *Aglais urticae* (Lepidoptera) based on DNA sequences of the mitochondrial COI gene and control region. *Molecular Phylogenetics and Evolution*, **31**, 630-646.

Vandewoestijne S, Nève G, Baguette M (1999) Spatial and temporal population genetic structure of the butterfly *Aglais urticae* (Lepidoptera, Nymphalidae). *Molecular Ecology*, **8**, 1539-1543.

Viñas J, Bremer JA, Pla C (2004) Phylogeography of the Atlantic Bonito (*Sarda sarda*) in the Northern Mediterranean: the combined effects of historical vicariance, population expansion, secondary invasion, and isolation by distance. *Molecular Phylogenetics and Evolution*, **33**, 32-42.

Wallwork J (1973) Zoogeography of some terrestrial micro-Arthropoda in Antarctica. *Biological Review*, **48**, 233-259.

Wang J-P, Lin H-U, Huang S, Pan C-H. Chen X-L, Chiang T-Y (2004) Phylogeography of *Varicorhinus barbatulus* (Cyprinidae) in Taiwan based on nucleotide variation of mtDNA and allozymes. *Molecular Phylogenetics and Evolution*, **31**, 1143-1156.

Wardle DA, Bardgett RD, Klironomos JN, Setälä H, Van der Putten WH, Wall DH (2004) Ecological linkages between aboveground and belowground biota. *Science*, **304**, 1629-1633.

Watson K (1967) The terrestrial Arthropoda of Macquarie Island. *ANARE Scientific reports* B99. p 90.

Whalley P, Jarzembowski EA (1981) A new assessment of *Rhyniella,* the earliest known insect, from the Devonian of Rhynie, Scotland. *Nature*, **291**, 317.

Whitlock MC, McCauley DE (1999) Indirect measures of gene flow and migration: $F_{ST} \neq 1/(4Nm+1)$. *Heredity*, **82**, 117-125.

Wiens JJ, Donoghue MJ (2004) Historical biogeography, ecology and species richness. *Trends in Ecology and Evolution*, **19**, 639-644.

Wolstenholme DR, Clary DO (1985) Sequence evolution of *Drosophila* mitochondrial DNA. *Genetics*, **109**, 725-744.

Wright S (1922) Coefficients of inbreeding and relationship. *American Naturalist*, **56**, 330-354.

Wright S (1931) Evolution in Mendelian populations. *Genetics*, **28**, 97-159.

Wright S (1943) Isolation by distance. *Genetics*, **28**, 114-138.

Wright S (1951) The genetical structure of populations. *Annals of Eugenetics,* **15**, 323-354.